This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

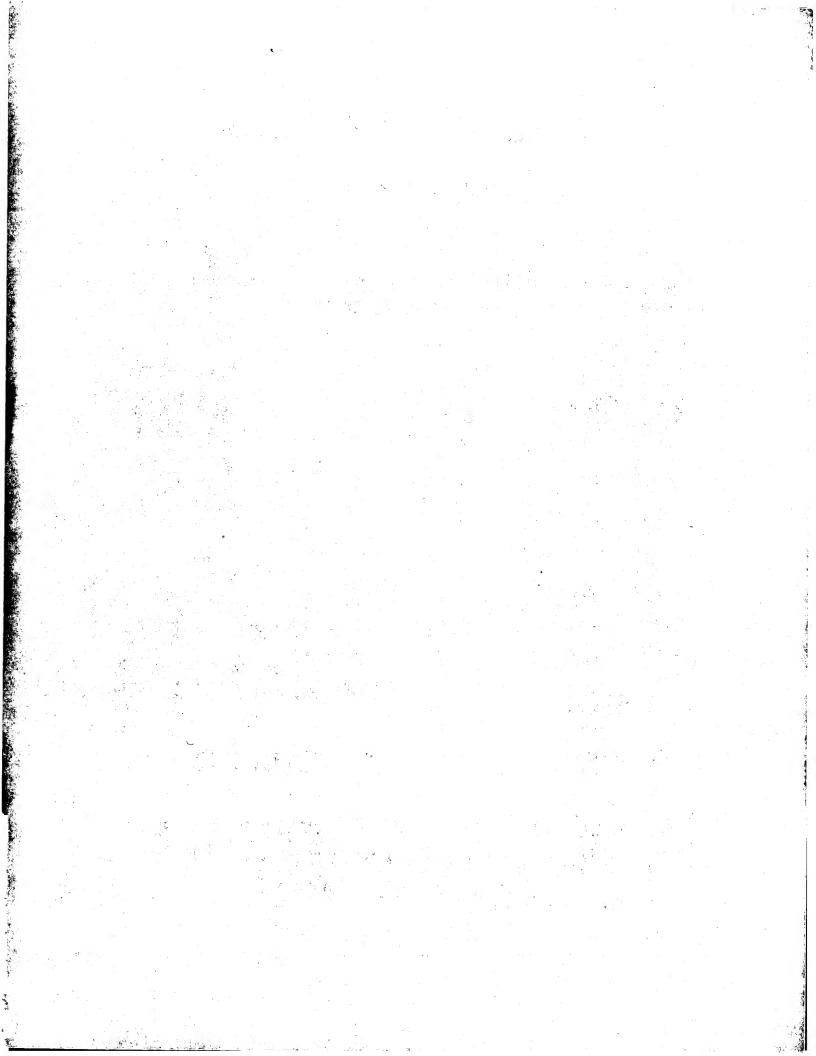
Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶:
C12N 15/29, C07K 14/415, C12N 5/04, 5/14, A01H 5/00, C12N 15/10, 15/82

(11) International Publication Number:

WO 99/19492

(43) International Publication Date:

22 April 1999 (22.04.99)

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

ΑU

(71) Applicant (for all designated States except US): RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): DOUTRIAUX, Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40-Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux, F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).
- (74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: METHODS FOR OBTAINING PLANT VARIETIES

(57) Abstract

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES .	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	Prance	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG:	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	Œ	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JР	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
СН	Switzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
СМ	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	ΚZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Methods for Obtaining Plant Varieties

TECHNICAL FIELD

The present invention relates to nucleotide sequences which encode polypeptides involved in the DNA mismatch repair systems of plants, and to the polypeptides encoded by those nucleotide sequences. The invention also relates to nucleotide sequences and polypeptide sequences for use in altering the DNA mismatch repair system in plants. The invention also relates to a process for altering the DNA mismatch repair system of a plant cell, to a process for increasing genetic variations in plants and to processes for obtaining plants having a desired characteristic.

BACKGROUND OF THE INVENTION

Plant breeding essentially relies on and makes use of genetic variation which occurs naturally within and between members of a family, a genus, a species or a subspecies. Another source of genetic variation is the introduction of genes from other organisms which may or may not be related to the host plant.

Allelic loci or non-allelic genes which constitute or contribute to desired quantitative (e.g. growth performance, yield, etc.) or qualitative (e.g. deposition, content and composition of seed storage products; pathogen resistance genes; etc.) traits that are absent, incomplete or inefficient in a species or subspecies of interest are typically introduced by the plant breeder from other species or subspecies, or *de novo*. This introduction is often done by crossing, provided that the species to be crossed are sexually compatible. Other means of introducing genomes, individual chromosomes or genes into plant cells or plants are well known in the art. They include cell fusion, chemically aided transfection (Schocher et al., 1986, Biotechnology 4: 1093) and ballistic (McCabe et al., 1988, Biotechnology 6: 923), microinjection (Neuhaus et al., 1987, TAG 75: 30), electroporation of protoplasts (Chupeau et al., 1989, Biotechnology 7: 53) or microbial transformation methods such as Agrobacterium mediated transformation (Horsch et al., 1985, Science 227: 1229; Hiei et al., 1996, Biotechnology 14: 745).

T.

However, when a foreign genome, chromosome or gene is introduced into a plant, it will often segregate in subsequent generations from the genome of the recipient plant or plant cell during mitotic and meiotic cell divisions and, in consequence, become lost from the host plant or plant cell into which it had been introduced. Occasionally, however, the introduced genome, chromosome or gene physically combines entirely or in part with the genome, chromosome or gene of the host plant or plant cell in a process which is called recombination.

Recombination involves the exchange of covalent linkages between DNA molecules in regions of identical or similar sequence. It is referred to here as homologous recombination if donor and recipient DNA are identical or nearly identical (at least 99%)

10

base sequence identity), and as homeologous recombination if donor and recipient DNA are not identical but are similar (less than 99% base sequence identity).

The ability of two genomes, chromosomes or genes to recombine is known to depend largely on the evolutionary relation between them and thus on the degree of sequence similarity between the two DNA molecules. Whereas homologous recombination is frequently observed during mitosis and meiosis, homeologous recombination is rarely or never seen.

From a breeder's perspective, the limits within which homologous recombination occurs, therefore, define a genetic barrier between species, varieties or lines, in contrast to homeologous recombination which can break this barrier. Homeologous recombination is thus of great importance for plant breeding. Accordingly there is a need for a process for enhancing the frequency of homeologous recombination in plants. In particular, there is a need for a process of increasing homeologous recombination to significantly shorten the length of breeding programs by reducing the number of crosses required to obtain an otherwise rare recombination event.

At least in Escherichia coli, homologous and homeologous recombination are known to share a common pathway that requires among others the proteins RecA, RecB, RecC. RecD and makes use of the SOS induced RuvA and RuvB, respectively. It has been suggested that mating induced recombination follows the Double-Strand Break Repair 20 model (Szostak et al., 1983, Cell 33, 25-35), which is widely used to describe genetic recombination in eukaryotes. Following the alignment of homologous or homeologous DNA double helices the RecA protein mediates an exchange of a single DNA strand from the donor helix to the aligned recipient DNA helix. The incoming strand screens the recipient helix for sequence complementarity, seeking to form a heteroduplex by hydrogen 25 bonding the complementary strand. The displaced homologous or homeologous strand of the recipient helix is guided into the donor helix where it base pairs with its counterpart strand to form a second heteroduplex. The resulting branch point then migrates along the aligned chromosomes thereby elongating and thus stabilising the initial heteroduplexes. Single stranded gaps (if present) are closed by DNA synthesis. The strand cross overs 30 (Holliday junction) are eventually resolved enzymatically to yield the recombination products.

Although in wild type E. coli homologous and homeologous recombination are thus mechanistically similar if not identical, homologous recombination in conjugational crosses E. coli x E. coli occurs five orders of magnitude more frequently than homeologous recombination in conjugational crosses E. coli x S. typhimurium (Matic et al. 1995; Cell 80, 507-515). The imbalance in favour of homologous recombination was shown to be caused largely by the bacterial MisMatch Repair (MMR) system since its

inactivation increased the frequency of homeologous recombination in *E. coli* up to 1000 fold (Rayssiguier et al. 1989, Nature 342, 396-401).

In E. coli. the MMR system (reviewed by Modrich 1991, Annual Rev Genetics 25, 229-253) is composed of only three proteins known as MutS, MutL and MutH. MutS recognizes and binds to base pair mismatches. MutL then forms a stable complex with mismatch bound MutS. This protein complex now activates the MutH intrinsic single stranded endonuclease which nicks the strand containing the misplaced base and thereby prepares the template for DNA repair enzymes.

During recombination, MMR components inhibit homeologous recombination. In vitro experiments demonstrated that MutS in complex with MutL binds to mismatches at the recombination branch point and physically blocks RecA mediated strand exchange and heteroduplex formation (Worth et al., 1994; PNAS 91, 3238-3241). Interestingly, the SOS dependent RuvAB mediated branch migration is insensitive to MutS/MutL, explaining the observed slight increase in SOS dependent homeologous recombination.

15 Homeologous mating even induces the SOS response, thereby taking advantage of RuvAB induction (Matic et al. 1995, Cell 80, 507-515).

The MMR system thus appears to be a genetic guardian over genome stability in E. coli. In this role it essentially determines the extent to which genetic isolation, that is, speciation, occurs. The diminished sensitivity of the SOS system to MMR, however, allows (within limits) for rapid genomic changes at times of stress, providing the means for fast adaptation to altered environmental conditions and thus contributing to intraspecies genetic variation and species evolution.

1

14.

The important role of MMR in preserving genomic integrity has been established also in certain eukaryotes. In its efficiency, the human MMR, for example, may even counteract potential gene therapy tools such as triple-helix forming oligonucleotides including RNA-DNA hybrid molecules (Havre et al., 1993, J. Virology 67: 7234-7331; Wang et al., 1995, Mol. Cell. Biol. 15: 1759-1768; Kotani et al., 1996, Mol. Gen. Genetics 250: 626-634; Cole-Strauss et al., 1996, Science 273: 1387-1389). Such oligonucleotides are designed to introduce single base changes into selected DNA target sequences in order to inactivate for example cancer genes or to restore their normal function. The resulting base mismatches however are recognised by the mismatch repair system which then directs removal of the mismatched base, thereby reducing the efficiency of oligonucleotide induced site-specific mutagenesis.

To date, two families of related genes, homologous to the bacterial MutS and MutL genes have been identified or isolated in yeast and mammals (recent reviews by Arnheim and Shibata, 1997, Curr. Opinion Genet. Dev. 7, 364-370; Modrich and Lahue, 1996, Annual Rev. Biochem. 65, 101-133; Umar and Kunkel, 1996, Eur. J. Biochem. 238, 297-307). Biochemical and genetic analysis indicated that eukaryotic MutS homologs (MSH)

WO 99/19492 4 PCT/EP98/06977

and MutL homologs (MLH, PMS), respectively, fulfil similar protein functions as their bacterial counterparts. Their relative abundance, however, could reflect different mismatch specificity and/or specialisation for different tissues or organelles or developmental processes such as mitotic versus meiotic recombination.

To date, six different genes homologous to *MutS* have been isolated in yeast (yMSH), and their homologs have been found in mouse (mMSH) and human (hMSH), respectively. Encoded proteins yMSH2, yMSH3 and yMSH6 appear to be the main MutS homologs involved in MMR during mitosis and meiosis in yeast, where the complementary proteins MSH3 and MSH6 alternatively associate with MSH2 to recognise different mismatch substrates (Masischky et al., 1996, Genes Dev. 10, 407-420). Similar protein interactions have been demonstrated for the human homologs hMSH2, hMSH3 and hMSH6 (Acharya et al., 1996, PNAS 93, 13629-13634).

MutL homologs (MLH and PMS), recently reviewed by Modrich and Lahue (1996, Annual Rev. Biochem. 65. 101-133) have so far been found in yeast (yMLH1 and yPMS1), mouse (mPMS2) and human (hMLH1, hPMS1 and hPMS2). The hPMS2 is a member of a family of at least 7 genes (Horii et al., 1994, Biochem. Biophys. Res. Commun. 204, 1257-1264) and its gene product is most closely related to yPMS1. Prolla et al. (1994, Science 265, 1091-1093) presented evidence for yPMS1 and yMLH1 to physically associate with each other and, together, to interact with the MutS homolog yMSH2 to form a ternary complex involved in mismatch substrate binding.

However, while medical interest in mismatch repair has prompted extensive research on MMR in bacteria, yeast and mammals, MMR genes have not been isolated from higher plants prior to the present invention and no attempts to adjust the plant MMR to plant breeding needs have been reported.

SUMMARY OF THE INVENTION

According to a first embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant. In one form of this embodiment, the invention provides an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human. More particularly, the invention provides polynucleotide sequences encoding polypeptides which are homologous to the mismatch repair polypeptides MSH3 and MSH6 of Saccharomyces cerevisiae. Still more particularly, the invention provides the coding sequences of the genes AtMSH3 and AtMSH6 of Arabidopsis thaliana. as defined hereinbelow, and polynucleotide sequences encoding polypeptides which are homologous to polypeptides encoded by AtMSH3 and AtMSH6.

25

ï

溢

···· 🎉

4.50

According to a second embodiment of the invention, there is provided an isolated and purified polypeptide functionally involved in the DNA mismatch repair system of a plant, for example a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human such as a polypeptide encoded by the genes AtMSH3 or AtMSH6 of 5 Arabidopsis thaliana, as defined hereinbelow.

According to a third embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

According to a fourth embodiment of the invention there is provided a chimeric gene comprising a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant: together with at least one regulation element capable of functioning in a plant cell. Examples of such regulation elements include constitutive, inducible, tissue type specific and cell type specific promoters such as 35S. NOS, PR1a. AoPR1 and DMC1. Typically, a chimeric gene of the fourth embodiment will also include at least one terminator sequence, more typically exactly one terminator sequence.

In the third and fourth embodiments, said interference, by said polynucleotide sequence, with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair peptide of a yeast or a human typically occurs by hybridisation or by co-suppression.

According to a fifth embodiment of the invention there is provided a plasmid or vector comprising a chimeric gene of the fourth embodiment. A vector of the fifth 30 embodiment may be, for example, a viral vector or a bacterial vector.

According to a sixth embodiment of the invention, there is provided a plant cell stably transformed, transfected or electroporated with a plasmid or vector of the fifth embodiment.

According to seventh embodiment of the invention, there is provided a plant as comprising a cell of the sixth embodiment.

According to an eighth embodiment of the invention, there is provided a process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising

WO 99/19492 PCT/EP98/06977

transforming or transfecting said plant cell with a DNA sequence of the third embodiment or a chimeric gene of the fourth embodiment or a plasmid or vector of the fifth embodiment, and causing said DNA sequence to express said polynucleotide or said polypeptide.

According to a ninth embodiment of the invention, there is provided a process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred. For example, homeologous recombination may be evidenced by new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait.

According to a tenth embodiment of the invention there is provided a process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.

In a preferred form of the ninth and tenth embodiments of the invention, the step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene of the fourth embodiment and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant or cells.

In other embodiments, the invention provides (a) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule of the first embodiment; (b) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO: 18 and (c) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30; with the proviso that the oligonucleotide of (a), (b) and (c) is other than SEQ ID NO:1 or SEQ ID NO:2.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a diagrammatic representation of the primer sequences used to 35 isolate AtMSH3.

Figure 2 is a plasmid map of clone 52, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH.

ier.

- 242

de

1

Figure 3 is a plasmid map of clone 13, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH3.

Figure 4 is a sequence listing of the coding sequence of AtMSH3, together with a deduced sequence of the encoded polypeptide.

Figure 5 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH3 protein.

Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6.

Figure 7 is a plasmid map of clone 43, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH6.

Figure 8 is a plasmid map of clone 62, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH6.

Figure 9 is a sequence listing of the coding sequence of AtMSH6, together with a deduced sequence of the encoded polypeptide.

Figure 10 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH6 protein.

Figure 11 is a genomic sequence listing of AtMSH6.

Figure 12 is a plasmid map of plasmid pPF13.

Figure 13 is a plasmid map of plasmid pPF14.

Figure 14 is a plasmid map of plasmid pCW186.

Figure 15 is a plasmid map of plasmid pCW187.

Figure 16 is a plasmid map of plasmid pPF66.

Figure 17 is a plasmid map of plasmid pPF57.

Figure 18 is a diagrammatic representation of an antisense gene construction for use in homeologous meiotic recombination.

Figure 19 is a plasmid map of plasmid p3243.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the inventors' discovery that there exist in higher plants genes which are homologous to MMR genes in E. coli, and to MMR genes in yeasts and humans.

Thus, the inventors have identified genes, herein designated AtMSH3 and AtMSH6, of the plant Arabidopsis thaliana which encode the proteins AtMSH3 and AtMSH6. These plant proteins are homologous to yMSH3 and yMSH6, respectively. The present inventors have isolated cDNAs encoding the proteins AtMSH3 and AtMSH6 and have isolated the complete gene encoding AtMSH6. Given the teaching herein, other genes (for example AtMSH2, and genes of other plants) may be obtained which are involved in DNA mismatch repair in plants, including other genes which encode polypeptides homologous to MMR proteins, of yeasts or humans, such as genes which encode

WO 99/19492 PCT/EP98/06977

polypeptides homologous to yeast MSH2. MLH1 or PMS2, or to human MLH1. PMS1 or PMS2. For example, given the teaching herein, genes of members of the *Brassicaceae* family or of other unrelated families, for example the *Poaceae*, the *Solanaceae*, the *Asteraceae*, the *Malvaceae*, the *Fabaceae*, the *Linaceae*, the *Canabinaceae*, the *Dauaceae* and the *Cucurbitaceae* family, and which encode polypeptides homologous to MMR proteins of yeasts or humans may be obtained.

Examples of plants whose genes encoding polypeptides homologous to MMR proteins of yeasts or humans may be obtained given the teaching herein include maize, wheat, oats, barley, rice, tomato, potato, tobacco, capsicum, sunflower, lettuce, artichoke, safflower, cotton, okra, beans of many kinds including soybean, peas, melon, squash, cucumber, oilseed rape, broccoli, cauliflower, cabbage, flax, hemp, hops and carrot.

Within the meaning of the present invention, a first polypeptide is defined as homologous to a second polypeptide if the amino acid sequence of the first polypeptide exhibits a similarity of at least 50% on the polypeptide level to the amino acid sequence of the second polypeptide.

A procedure which may be followed to obtain genes AtMSH3 and AtMSH6 is described in Example 1. Essentially the same technique may be applied to obtain other mismatch repair genes of Arabidopsis thaliana, and essentially the same technique as 20 exemplified herein may be applied to cDNA obtained by reverse transcription of RNA from other plants. Alternatively, given the sequence information disclosed herein, other degenerate oligonucleotide primers, especially oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions (such as the conditions described in Example 1 using primers UPMU and DOMU) to AtMSH3 and/or AtMSH6 25 may be designed and obtained for use in isolating sequences of plant mismatch repair genes which are homologous to AtMSH3 or AtMSH6, from other plants. oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions to plant mismatch repair genes of plants other than Arabidopsis thaliana also fall within the scope of the present invention and may be utilised to obtain mismatch Typically, such oligonucleotides are capable of 30 repair genes of still other plants. hybridising at 45°C under standard PCR conditions to a DNA molecule which encodes a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or a human. The temperature at which oligonucleotides of the invention hybridise to AtMSH3 and/or AtMSH6, or to plant mismatch repair genes of plants other than Arabidopsis thaliana, or 35 to DNA molecules which encode polypeptides which are homologous to a mismatch repair polypeptide of a yeast or a human may be higher than 45°C, for example at least 50°C, or 6 at least 55°C, or at least 60°C or as high as 65°C.

The successful gene isolation disclosed herein demonstrates for the first time the existence of MMR in higher plants and indicates the presence of other plant MMR genes. For example, genes encoding the plant homologs of MSH1, MSH2, MSH4, MSH5, PMS1, PMS2 and MLH1 may be identified given the teaching herein. Such genes, as well as those specifically described herein, separately or in combination, are useful in manipulating the plant MMR for plant breeding purposes. Thus, for example, the plant MMR may be altered by including in a plant cell a polynucleotide sequence as defined herein above with reference to the third embodiment of the invention, and causing the polynucleotide sequence to express either a polynucleotide which disables a plant MMR gene, or a polypeptide which disrupts the plant's MMR system.

The DNA molecule of the third embodiment of the invention includes a polynucleotide sequence (herein referred to as a MMR altering gene) which may for example encode sense, antisense or ribozyme molecules characterised by sufficient base sequence similarity or complementarity to the gene to be altered to permit the antisense or 15 ribozyme molecule to hybridise with the plant MMR gene in vivo or to permit the sense molecule to participate in co-suppression. Alternatively, the MMR altering gene may encode a protein or proteins which interfere with the activity of a plant MMR protein and thus disrupt the plant's MMR system. For example, such encoded proteins may be antibodies or other proteins capable of interfering with MMR protein function, such as by 20 complexing with a protein functionally involved in plant MMR thereby disrupting the MMR of the plant. An example of such a protein is the MSH3 protein of Arabidopsis thaliana described herein or a protein of another plant which is homologous to the MSH3 protein of A. thaliana. For instance, overexpression of MSH3 in a plant cell causes MSH2 present in the cell to be substantially completely complexed, disrupting the 25 mismatch repair mechanism or mechanisms in the cell which are functionally dependent on the presence of a complex of MSH2 with MSH6. Similarly, mismatch repair mechanisms which depend on the presence of a complex of MSH2 and MSH3 may be disrupted by the overexpression of MSH6.

- : :

- 1,

A chimeric gene of the fourth embodiment, incorporating a MMR altering gene, may be prepared by methods which are known in the art. Similarly, the MMR altering gene may be introduced into a plant cell, regenerating tissue or whole plant by techniques known in the art as being suitable for plant transformation, or by crossing. Known transformation techniques include Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters. Suitable promoters may direct constitutive expression, such as the 35S or the NOS promoter. Usually, however, the promoter will direct either inducible or tissue specific (e.g. callus; embryonic (issue; etc.), cell type specific (e.g. protoplasts; meiocytes; etc.) or developmental (e.g. embryo) expression of the altering gene or genes, in order for the

MMR system to function in tissue types or cell types, or at developmental stages of the plant, in which it is not desirable for the MMR system to be altered. Using such promoters, therefore, the activity of a MMR altering gene may be limited to a specific stage during plant development or it may be altered by controlling conditions external to the plant, and the deleterious effects of a permanently disabled or altered DNA mismatch repair system in a plant may be avoided. Examples of suitable promoters which are not constitutive are known in the art and include inducible promoters such as PR1a (reviewed by Gatz, 1997, Annual Rev. Plant Phys. Plant Mol. Biol. 48: 89), tissue specific promoters such as AoPR1 (Sabahattin et al., 1993, Biotechnology 11: 218), and cell-type specific promoters such as DMC1.

A chimeric gene in accordance with the invention may further be physically linked to one or more selection markers such as genes which confer phenotypic traits such as herbicide resistance, antibiotic resistance or disease resistance, or which confer some other recognisable trait such as male sterility, male fertility, grain size, colour, growth rate, flowering time, ripening time, etc.

The process of the tenth embodiment of the invention provides, for example, a process for generating intraspecies genetic variation by altering the mismatch repair system in a plant cell, in regenerating plant tissue or in a whole plant. The plant cell, regenerating tissue or whole plant includes and expresses one or more MMR altering 20 genes which are capable of altering mismatch repair in the plant cell, regenerating tissue or whole plant. Alteration of MMR may be achieved, for example, by inactivating the genes encoding plant MSH3 and/or plant MSH6. It is preferred to inactivate the plant MSH3 and MSH6 encoding genes at the same time and in the same plant cell, regenerating tissue or whole plant. Typically in this preferred form of the invention 25 inactivation of either plant MSH3 or MSH6 alone is insufficient to substantially alter the plant's mismatch repair system and only when both MSH3 and MSH6 are inactivated simultaneously is the plant's mismatch repair system sufficiently altered to prevent the MMR system from recognising base pair mismatches, base insertions or deletions as a result of DNA replication errors. DNA damage, or oligonucleotide induced site-specific 30 mutagenesis. However, in some applications of the invention, inactivation of only one gene may also be used to cause genomic instability or increase the efficiency of sitespecific mutagenesis.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the genome of the plant cell, regenerating tissue or whole plant in order to restore mismatch repair in the plant cell, regenerating tissue or whole plant. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools, such as ribozymes, or may be removed from the genome using gene elimination systems known in the art, such as CRE/LOX. It is preferred to render two genes, whose gene products combine to incapacitate MMR, ineffective by separating

the altering genes through segregation. Therefore, in a preferred embodiment of the invention a first plant cell or plant is generated in which only plant MSH3 is incapacitated, and a second plant cell or plant is generated in which only plant MSH6 is incapacitated. The combination of both genomes, for example by crossing, then produces significant MMR deficiency in those cells or plants which have inherited both altering genes. If the altering genes are expressed from unlinked loci, gene segregation restores MMR activity in the progeny of the cells or plants.

In a process of the ninth embodiment of this invention, homeologous recombination is enhanced between different genomes, chromosomes or genes in plant cells or plants by altering MMR in said plant cells or plants. Such genomes, chromosomes or genes are characterised in that they originate from different plant families, genera, species, subspecies, plant varieties or lines. Hybrid plant cells or hybrid plants may be produced by crossing, by cell fusion or by other techniques known in the art. These plant cells or plants are further characterised by expressing one or more genes that are capable of altering mismatch repair in the plant cell or plants.

In the process of the ninth embodiment, the homeologous recombination is typically for the purpose of introducing a desired characteristic in the hybrid plant. In this typical application of the process of the ninth embodiment, and in the process of the tenth embodiment the desired characteristic may be any characteristic which is of value to the plant breeder. Examples of such characteristics are well known in the art and include altered composition or quality of leaf or seed derived storage products (e.g. oil, starch, protein), altered composition or quality of cell walls (e.g. decrease in lignin content), altered growth rate, prolonged flowering, increased plant yield or grain yield, altered plant morphology, resistence to pathogens, tolerance to or improved performance under environmental stresses of various kinds, etc.

In a preferred form of the tenth embodiment, an MMR altering gene is cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant to produce a
hybrid plant cell or hybrid plant in which homeologous recombination can occur.

Suitably, the MMR proficient plant cell or MMR proficient plant may also include an
MMR altering gene. For example a gene capable of inactivating plant MSH3 may be cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant in which MSH6
is inactivated. A resultant hybrid plant in which homeologous recombination occurs will
include both the MSH3 and MSH6 altering genes and its MMR system will therefore be
inactivated.

In this form of the invention, if hybrid plants are to be produced by crossing, the MMR altering gene preferably originates from the male parent, thus ensuring that the

10.

MMR altering gene is always introduced and is not present in the recipient cell. That is, the MMR of the recipient cell, prior to introduction of the MMR altering gene, is typically proficient. Alternatively, if an MMR altering gene is present in a recipient cell it may be ineffective or inefficient on its own, or it may be linked to an inducible or tissue specific or cell type specific promoter which only renders the MMR altering gene active under limited conditions.

Thus, in a preferred form of the process of the ninth embodiment, the MMR system of the hybrid plant is initially unaltered. In this form of the process, the step of altering the mismatch repair system may comprise introducing into the hybrid plant, or cells thereof, a MMR altering gene, such as by Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters, as described above. Preferably, the promoter is transcriptionally active in mitotically and meiotically active tissue and/or cells to ensure MMR alteration after chromosome pairing at mitosis and meiosis, respectively. The preferred timing for MMR alteration is at meiosis, because recombinant genomes, chromosomes or genes are directly transmitted to the progeny. A suitable meiocyte specific promoter is for example the DMC1 promoter from Arabidopsis thaliana ssp. Ler. (Klimyuk and Jones, 1997, Plant J. 11, 1-14). However, mitotic homeologous recombination is also a desirable outcome as somatic recombination events can be transmitted to offspring due to the totipotency of plant cells and the lack of predetermined germ cells in plants.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the hybrid plant or hybrid plant cells, in order to restore mismatch repair in the hybrid plant or hybrid plant cells. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools as described herein above.

EXAMPLES

Example 1. Cloning of the AtMSH3 and AtMSH6 coding sequences

Isolation of partial AtMSH3 and AtMSH6 consensus sequences

Degenerate oligonucleotides UPMU (SEQ ID NO:1) and DOMU (SEQ ID NO:2)

UPMU CTGGATCCACIGGICCIAA(C/T)ATG

DOMU CTGGATCC(A/G)TA(A/G)TGIGTI(A/G)C(A/G)AA

were used to isolate AtMSH3 and AtMSH6 sequences by PCR amplification.

Primers UPMU and DOMU correspond to conserved amino acid sequences of the proteins MutS (E. coli and S. typhimurium), HexA (S. pneumoniae). Repl (mouse) and Ducl (human). The conserved regions to which they are targeted are TGPNM for UPMU (amino acid positions 852-856 for AtMSH6 and 816-820 for AtMSH3) FATHY or FVTHY

30

. 透

· 🛬

for DOMU (amino acid positions 964-968 for AtMSH6 and 928-932 for AtMSH3, respectively.) These primers have been used to isolate MSH2 and MSH1 from yeast (Reenan and Kolodner, Genetics 132: 963-973 (1992)) and MSH2 from *Xenopus* and mouse (Varlet et al., Nuc. Acids Res. 22:5723-5728 (1994)).

Template single strand cDNA was produced by reverse transcription of 2 µg total RNA from a cell suspension culture of Arabidopsis thaliana ecotype Columbia (Axelos et al. 1989, Mol. Gen. Genetics 219: 106-112). The PCR reaction was performed under the following conditions in a final volume of 100µl: 0.2mM dNTP, 1µM each primer, 1XPCR buffer, 1u Taq DNA polymerase (Appligene) in the presence of template cDNA. PCR 10 parameters were 5 minutes at 94°C, followed by 30 cycles of 40 seconds at 95°C, 90 seconds at 45°C, 1 minute at 72°C. The amplification products were cloned into pGEM-T vector (Promega) and sequenced. Two different clones were isolated, S5 (350bp) was homologous to MSH3, S8 (327bp) was homologous to MSH6. Complete cDNA sequences were then isolated according to the Marathon cDNA amplification kit procedure (Clontech). 15 In summary, this procedure involves producing double stranded cDNA by reverse transcription of 2µg polyA+ RNA from the cell suspension culture of Arabidopsis, Adaptors are ligated on each side of the cDNA. The ligated cDNA is used as a template for 5° and 3° RACE PCR reactions in the presence of primers that are specific for the adaptor on one side (AP1 and AP2), and specific for the targeted gene on the other side. A 5, and a 3 20 fragment that overlap are thus produced for each gene. The complete gene coding sequence can be reconstituted taking advantage of a unique restriction site, if available, in the overlapping region. Specific details of this procedure as it was used to isolate AtMSH3 and AIMSH6 coding regions, are as follows.

Isolation of AtMSH3 complete coding sequence

From the sequence of clone S5, primer 636 (SEQ ID NO:3) was designed:

TGCTAGTGCCTCTTGCAAGCTCAT.

Primer AP1 (SEQ ID NO:4) is complementary to a portion of an adaptor sequence which had been ligated to the 5' and 3' ends of Arabidopsis cDNA:

API CCATCCTAATACGACTCACTATAGGGC.

PCR performed on the ligated cDNA with primers 636 and AP1 for the 5' RACE PCR was followed by a second round of amplification with the nested primers AP2 (SEQ ID NO:5) and S525 (SEQ ID NO:6)

AP2 ACTCACTATAGGGCTCGAGCGGC

S525 AGGTTCTGATTATGTGTGACGCTTTACTTA

35 (the latter was also designed to correspond to a part of the sequence of clone S5) and produced a 2720bp DNA fragment. Figure 1 provides a diagrammatic representation of the primer sequences used to isolate AtMSH3. Another primer (S51, SEQ ID NO:7)

7

S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG

25

was designed closer to the 5' border and permitted the determination of 99bp upstream to the ATG initiation codon. For the 3' RACE PCR, a first PCR reaction was performed with primers AP1 and 635 (SEQ ID NO:8).

635 GCACGTGCTTGATGGTGTTTTCAC

5 followed by a second round of amplification, using the nested primers AP2 and S523 (SEQ ID NO:9)

S523 TCAGACAGTATCCAGCATGGCAGAAGTA

which produced a DNA fragment of 890bp. Both DNA fragments were subcloned into pGEM-T and sequenced. Since PCR amplification using the Expand Long Template PCR to System (Boehringer-Mannheim) produced errors in the sequence, new oligonucleotides were designed to isolate those sequences again by PCR, but with the high fidelity DNA polymerase Pfu. PCR with primers 1S5 (SEQ ID NO:10) and S53 (SEQ ID NO:11)

- 1S5 ATCCCGGGATGGGCAAGCAAAAGCAGCAGACGA
- S53 GACAAAGAGCGAAATGAGGCCCCTTGG
- amplified the 1244bp fragment clone 52 (SEQ ID NO:12, cloned into pUC18/Sma1). PCR with primers S52 (SEQ ID NO:13) and 2S5 (SEQ ID NO:14)
 - 2S5 ATCCCGGGTCAAAATGAACAAGTTGGTTTTAGTC
 - S52 GCCACATCTGACTGTTCAAGCCCTCGC

amplified the 2104bp clone 13 (SEQ ID NO:15, cloned into pUC18/Sma1). The complete coding sequence of the AtMSH3 gene was reconstructed in pUC18 by ligating the 5' half of AtMSH3 (clone 52) to the 3' half of AtMSH3 (clone 13) after digesting with BamH1 which has a unique cleavage site in the overlapping region of both clones. This manipulation yielded plasmid pPF26. The Smal fragment from pPF26 contains the complete AtMSH3 coding sequence. The remaining primers referred to in Figure 1 are as follows:

- S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG (SEQ ID NO:16)
- S525 AGGTTCTGATTATGTGTGACGCTTTACTTA (SEQ ID NO:17)

Figures 2 and 3 provide plasmid maps of clones 52 and 13 respectively, showing restriction enzyme cleavage sites. The complete AtMSH3 coding sequence (SEQ ID NO:18) 30 is 3246bp long and is shown in Figure 4 together with the deduced sequence (SEQ ID NO:19) of the encoded polypeptide. AtMSH3 is clearly homologous to the yeast and mouse MSH3 genes. A sequence alignment of polypeptides encoded by AtMSH3 and that encoded by Saccharomyces cerevisiae MSH3 is set out in Figure 5.

Isolation of the AtMSH6 complete coding sequence and genomic sequences

The same procedure allowed isolation of the AtMSH6 cDNA. Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6. For the 5' RACE PCR, primers 638 (SEQ ID NO:20) and AP1 (SEQ ID NO:4)

638 TCTCTACCAGGTGACGAAAAACCG
allowed the amplification of a 2889 DNA fragment. Primer S81 (SEQ ID NO:21)

200

4,00

S81 CGTCGCCTTTAGCATCCCCTTCCTTCAC

helped define the 142bp upstream to the ATG initiation codon. On the 3' side, RACE PCR was initially performed with primers S823 (SEQ ID NO:22) and API (SEQ ID NO:4).

S823 GCTTGGCGCATCTAATAGAATCATGACAGG

5 and then with the nested primers 637 (SEQ ID NO:23) and AP2 (SEQ ID NO:5).

637 GACAGCGTCAGTTCTTCAGAATGC

to produce a 774bp DNA fragment. As for AtMSH3, those fragments were cloned and sequenced. Re-isolation of the DNA sequence using the high fidelity Pfu polymerase and newly designed primers 1S8 (SEQ ID NO:24) and S83 (SEQ ID NO:25) (for the 5' side) led to a 2182 bp DNA fragment identified as clone 43 (SEQ ID NO:26, cloned in pUC18/Sma1), and a 1379bp clone identified as clone 62 (SEQ ID NO:27, also cloned in pUC18/Sma1).

158	ATCCCGGGATGCAGCGCCAGAGATCGATTTTGT	
2S8	ATCCCGGGTTATTTGGGAACACAGTAAGAGGATT (SEQ	ID
NO:28)		
S82	GCGTTCGATCATCAGCCTCTGTGTTGC (SEQ ID NO:29)	•
S83	CGCTATCTATGGCTGCTTCGAATTGAG	

Figures 7 and 8 provide plasmid maps of clones 43 and 62 respectively, showing restriction enzyme cleavage sites. Clones 43 and 62 were digested by the Xmn1 restriction enzyme for which a unique site is present in their overlapping region and then ligated. The complete AtMSH6 coding sequence (SEQ ID NO:30) is 3330bp long and is shown in Figure 9 together with the deduced sequence (SEQ ID NO:31) of the encoded polypeptide. AtMSH6 is clearly homologous to the yeast and mouse MSH6genes. A sequence alignment of polypeptides encoded by AtMSH6 and that encoded by Saccharomyces cerevisiae MSH6 is set out in Figure 10.

An AtMSH6 genomic sequence was also isolated from a genomic DNA library constituted after partial Sau3AI digestion of DNA from the Arabidopsis cell suspension. 8062bp were sequenced that covered the AtMSH6 gene and show colinearity with the cDNA. 16 introns are found scattered along the gene. The complete genomic sequence 30 (SEQ ID NO:98) is shown in Figure 11.

Example 2. A measure of somatic variation in MMR deficient plants

Constructs

15

Constructs with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 under the control of a single 35S promoter have been inserted into the binary vector pPZP121 (Hajdukiewicz et al., Plant Mol. Biol. 23, 793-799) between the right and left borders of the T-DNA. The pPZP121 plasmid confers chloramphenicol resistance to Escherichia coli or Agrobacterium tumefaciens bacteria. The aacC1 gene is carried by the T-DNA and allows selection of transformed plant cells on gentamycin (Hajdukiewicz et al., Plant Mol. Biol. 25, 989-994). For the purpose of expressing antisense constructs, a 35S

promoter/terminator cassette with a polylinker was introduced into pPZP121. The 3' ends of the considered genes have been chosen since this region seems more efficient for antisense inhibition. For AtMSH3 this corresponds to clone 13 (2104bp), for AtMSH6 this corresponds to clone 62 (1379bp). Clone 13 comprises 2104bp of the 3' region that were cut off the pUC18 vector by Sal1/Sst1 restriction, blunted with T4 DNA polymerase and ligated into the T4 DNA polymerase blunted BamHI site of pPZP121/35S, creating clone pPF13. The same procedure was followed for the 3' region of AtMSH6 clone 62 (1379bp) thus creating plasmid pPF14. For the double constructs, the 3' region (from clone 62) of AtMSH6 was introduced ahead of the AtMSH3 region into pPF13 creating pCW186 and reciprocally, the 3' region of AtMSH3 (from clone 13) was introduced ahead of AtMSH6 into pPF14, creating pCW187.

These constructs were introduced into the Arabidopsis cells (as described below) of wildtype Columbia and of the Columbia tester line.

An alternative strategy to antisense inhibition of AtMSH6 comes from experiments of Marra et al. (1998, Proc. Natl. Acad. Sci USA 95, 8568-8573) who show that overexpression of functional MSH3 results in depletion of MSH6 protein in human cells. This depletion may generate a mismatch repair mutant phenotype.

For the purpose of overexpressing functional AtMSH3 protein in plant cells, the complete MSH3 coding region was excised from pPF26 (example 1) by digestion with 20 Smal, and was inserted into the Smal site of pCW164. The resulting construct was named pPF66. It contains a complete AtMSH3 gene under the control of the 35S promoter inside the left (LB) and right (RB) border of the T-DNA. This T-DNA also contains the hpt2 gene for gentamycin selection. Plasmid pPF66 was introduced into Arabidopsis cells as described below. One cell clone was selected which clearly overexpressed the AtMSH3 gene as shown by Northern analysis. Figures 12-16 provide plasmid maps of plasmids pPF13, pPF14, pCW186, pCW187 and pPF66 respectively.

Construction of tester construct

For the purpose of Forward Mutagenesis Assays, a tester construct was built containing the coding regions for nptII, codA, uidA. All three genes are driven by the 35S promoter and are terminated by the 35S terminator. This construct was obtained by introducing an EcoR1 fragment encoding the codA cassette (2.5kb) and a HindIII fragment encoding the uidA (GUS) cassette (2.4kb) into the pPZP111 vector (Hajdukiewicz et al.,1994, Plant Mol Biol 23: 793-799) which already contained the nptII expression cassette. This new plasmid was named pPF57. NptII is used to select for transformed plant cells.

35 GUS is used to analyse the degree of gene silencing in the construct (i.e. to identify cell lines in which the transgenes are expressed), and codA is used as a marker for forward mutagenesis (described below).

٠.

μ Έ. μ≟ ₁₄

The plasmid map of pPF57 is provided in Figure 17.

Plant cell transformation

then transformed by co-cultivation. A suspension culture of Arabidopsis thaliana cells that has been established by Axelos et al. (1992, Plant Physiol. Biochem. 30, 1-6) may be used. One day old freshly subcultured cells are diluted five times into AT medium (Gamborg B5 medium. 30g/l sucrose, 200µg/l NAA). 10µl of saturated Agrobacterium containing the transforming T-DNA constructs are added to 10ml diluted cells in a 100ml erlenmeyer. The co-cultivation is agitated slowly (80rpm) for 2 days. The cells are then washed 3 times into AT medium and finally resuspended in the same initial volume (10ml). The culture is agitated for 3 days to allow expression before plating on selection plates (AT/BactoAgar 8g/l+gentamycin 50µg/ml). Transformed individual calli are isolated 3 weeks later.

Tester Strain

The tester construct on plasmid pPF57 was introduced into Arabidopsis cells of wildtype Columbia using the transformation protocol described above. Among 10 candidate transformants, one cell clone was shown (by Southern analysis) to have a unique T2DNA insertion. All three genes were shown to be functional in this cell line as indicated by resistance to kanamycin, blue staining in the presence of X-Glu (GUS), and sensitivity to 5-fluoro-cytosine (codA).

MMR altering genes (described above) were then introduced individually into the tester line and transformed cells are used for analysis of both Microsatellite Instability and Forward Mutagenesis.

Microsatellite analysis

Microsatellites have been described in Arabidopsis (Bell and Ecker, 1994, Genomics 19, 137-144). The present Example is based on a study of instability of microsatellites that are polymorphic amongst different ecotypes. DNA is extracted from the transformed calli. Specific primers have been defined that are used to amplify the microsatellite sequence. One of the two primers is previously P³² labelled by T4 kinase. In case of a polymorphic variation, new PCR products appear that do not follow the expected pattern of migration on a polyacrylamide gel. This is a commonly observed feature for MMR deficient cells in yeast or mammalian cells.

In particular, the present Example describes a study on microsatellites ca72 (CA₁₈), nga172 (GA₂₉), and ATHGENEA(A₃₉), chosen because they belong to the types predominantly affected in human mismatch repair deficient tumors. The size of these microsatellites is not conserved from one *Arabidopsis* ecotype to the other.

Arabidopsis cells which are transformed with an MMR altering gene (above) and control cells not expressing the MMR altering gene are allowed to form calli. DNA is

rapidly extracted from the calli and is analysed for microsatellite instability as described in detail by Bell and Ecker 1994. Genomics 19, 137-144. In summary, the relevant microsatellite is amplified by PCR using P32 labelled primers. The PCR products are separated on a DNA sequencing gel for size determination. Size differences between microsatellites from transformed and control cells not expressing the MMR altering gene in question indicate microsatellite instability as a result of MMR alteration.

The sequences of primers used for PCR amplification of microsatellites ca72 and nga172 are included in Table 1. PCR amplification of microsatellite ATHGENEA made use of a forward primer containing the sequence

ACCATGCATAGCTTAAACTTCTTG (SEQ ID NO:32)

and of a reverse primer containing the sequence

ACATAACCACAAATAGGGGTGC (SEQ ID NO:33).

The amplification for microsatellite ca72 revealed in Columbia control cells (with respect to the MMR altering gene) a 248 bp long PCR fragment instead of the published length of 124 bp. DNA sequencing verified this fragment as a CA₁₈ microsatellite.

Forward mutagenesis assav

10

Tester cells transformed with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 are analysed for the stability of the codA gene. The functional codA gene confers to sensitivity to 5-fluoro-cytosine (5FC), whereas a gene inactivating mutation in 20 codA will confer resistance to 5FC. The frequency of resistant cells is therefore a good indicator of somatic variation as a direct result of MMR alteration. Variants resistant to 5FC are first analysed for GUS activity. If GUS is inactive, 5FC resistance is assumed to be due to gene silencing (all three genes are under the 35S promoter). If GUS is active, 5FC resistance is assumed to be due to forward mutations that have inactivated codA. PCR is then performed on the putative codA mutant genes which is then sequenced to confirm the presence of forward mutations in codA.

Besides codA, other marker genes may also be used for the Forward Mutagenesis Assay such as the ALS gene (conferring sensitivity to valine or to sulfonylurea; Hervieu and Vaucheret, 1996, Mol. Gen. Genet. 251 220-224; Mazur et al. 1987, Plant Physiol. 85 1110-30 1117).

Example 3. Homeologous meiotic recombination in Arabidopsis thaliana

A. Construction of a meiocyte specific gene expression cassette comprising the DMC1 promoter and the NOS terminator

(i) The DMC1 promoter may be used as published by Klimyuk and Jones, 1997, 35 Plant J. 11.1-14). To obtain a more convenient alternative for gene cloning, a 3.3 Kb

ý

long subfragment of the DMC1 promoter was obtained by PCR from genomic DNA of Arabidopsis thaliana (ssp. Landsberg erecta "Ler").

The PCR was done in three rounds:

Round One: A 3.7 Kb long product was obtained using the forward primer 5 DMCIN-A comprising the sequence

GAAGCGATATTGTTCGTG (SEQ ID NO:34)

and the reverse primer DMCIN-B comprising the sequence

AGATTGCGAGAACATTCC (SEQ ID NO:35).

The weak amplification product was then used as template for round two and three.

Round Two: A 3.1 Kb long product comprising the promoter and the 5' untranslated leader was obtained using forward primer DMCIN-1, which contained the sequence

acgcgtcgacTCAGCTATGAGATTACTCGTG (SEQ ID NO:36)

and introduced a Sall cloning site at the 5' end of the promoter fragment, and reverse primer DMCIN-2 which contained the sequence

gctctagaTTTCTCGCTCTAAGACTCTCT (SEQ ID NO:37)

and introduced a XbaI site at the 3' end of the PCR fragment.

Round Three: A 0.2 Kb long product comprising the first exon/intron of the *DMC1* promoter was obtained using forward primer DMCIN-3, which contained the sequence

gctctagaGCTTCTCTTAAGTAAGTGATTGAT (SEQ ID NO:38)

and introduced a XbaI site at the 5' end of the PCR fragment, and reverse primer DMCIN-4, containing the sequence

teccegggetegagagatetecatggTTTCTTCAGCTCTATGAATCC (SEQ ID NO:39) and introduced at the 3' end of the PCR product restriction sites for Ncol. Bg/II. Xhol and 25 Smal.

The products obtained in round Two and Three were digested with XbaI and subsequently ligated to reconstitute a 3.3 Kb long DMC1 promoter from which the first two in-frame ATG start codons were replaced with a unique restriction site for XbaI. This promoter can be cloned between the restriction sites for SaII and SmaI of p3264, which contains the SacI-EcoRI NOS terminator in pBIN19, to yield the entire expression cassette in pBIN19. This cassette contains the following cloning sites: NcoI, BgIII, XhoI. SmaI and (already present on p3264) KpnI and SacI.

(ii) Another strategy yielded the following convenient *DMC*1 promoter. A 1.8 kb DNA fragment comprising the 3' terminal part of the meiocyte specific *DMC*1 promoter was isolated by PCR from purified genomic DNA of *Arabidopsis thaliana* (ssp. Landsberg erecta "Ler"). The forward PCR primer (DMC1a) contained the sequence

acgcgtcgacGAATTCGCAAGTGGGG (SEQ ID NO:40)

and introduced a Sall cloning site at the 5' end of the promoter fragment. The reverse PCR primer (DMC1b) contained the sequence

20

introducing a polylinker region at the 3' end of the promoter fragment. The PCR reaction was carried out with VENT DNA Polymerase (NEB) over 25 cycles using the following

The PCR reaction yielded a blunt ended DNA fragment which was digested with restriction enzyme Sall and was cloned into the cleavage sites of restriction enzymes Sall and Smal in plasmid p2030, a pUC118 derivative containing the SacI-EcoRI NOS terminator fragment of pBIN121. The cloning yielded plasmid p2031, containing the DMC1 promoter-polylinker-NOS terminator expression cassette depicted in Figure 18.

cycling protocol: 1 minute at 94°C, 1 minute at 54°C, 2 minutes at 72°C.

B. Construction of an MSH3 antisense gene under the control of the DMC1 promoter

A 2.1 kb DNA fragment encoding the carboxyterminal part of AtMSH3 was removed from the polylinker of clone 13 described in Example 1 after (i) digestion with KpnI. (ii) blunting of the DNA ends generated by KpnI and (iii) digestion with BamHI. The isolated fragment was then cloned in antisense orientation downstream of the DMC1 promoter in plasmid p2031, which had been digested with restriction enzymes SmaI and Bg/II. This cloning yielded plasmid p2033 (Figure 18).

After digestion of p2033 with EcoRI, a 4.1 kb DNA fragment was recovered comprising the DMC1 promoter, the partial MSH3 cDNA in antisense orientation with respect to the promoter and the NOS terminator. This fragment was cloned into the EcoRI restriction site of plant transformation vector pNOS-Hyg-SCV to yield plasmid p3242 (Figure 18).

C. Construction of a combined MSH6/MSH3 antisense gene under the control of a single DMC1 promoter

A 3.1 kb fragment, encoding in antisense orientation the partial AtMSH6 and AtMSH3 sequences and the 35S terminator, was isolated from pCW186 by digestion with KpnI. The fragment was treated with Klenow enzyme to blunt both ends. It was then cloned into the SmaI site of plasmid p3243 (a pNOS-Hyg-SCV derivative, illustrated in Figure 19), which had been opened to delete the region between the SmaI sites. Clones containing the fragment in the antisense orientation with respect to the DMC1 promoter (described in A(ii) above) were identified by diagnostic digestion with BamHI. The selected construct was named p3261.

Another practical way of cloning the double antisense gene is as follows. A 1 kb DNA fragment encoding the carboxyterminal part of AtMSH6 is isolated from clone 62 described in Example 1 after digestion of clone 62 plasmid DNA with BamHI, which cleaves in the 5' polylinker region flanking the partial cDNA, and with EcoRI, which cleaves within the cDNA. The isolated fragment is treated with Klenow enzyme to blunt both its ends and is cloned into the recipient plasmid p2033 or p3242. For the purpose of

cloning, the recipient plasmid may be cleaved with either AvaI or NcoI and can be blunted with Klenow enzyme to produce blunt acceptor ends for fragment cloning. This cloning yields two opposite orientations of cloned fragment DNA with respect to the DMC1 promoter. These can be identified by diagnostic digestion with restriction enzymes ScaI or XmnI in conjunction with SacI. The selected construct contains the DMC1 promoter, the combined partial cDNAs for AtMSH3 and AtMSH6 (both cloned in antisense orientation with respect to the DMC1 promoter) and the NOS terminator. If the recipient plasmid is p2033, the combined antisense gene under control the single DMC1 promoter is recovered from the vector after EcoRI digestion and is cloned into the EcoRI restriction site of pNOS-Hyg-SCV.

D. Construction of a full-length MSH3 sense gene under control of the DMC1 promoter for overexpression of functional MSH3 protein

Overexpression of MSH3 protein was shown in human cells (Marra et al., 1998, Proc. Natl. Acad. Sci. USA 95, 8568-8573) to complex all available MSH2 protein. This leaves MSH6 protein without its partner, leading to the degradation of MSH6 protein and eventually to a mismatch repair phenotype.

This phenomenon is exploited to increase homeologous meiotic recombination in Arabidopsis as an alternative to antisense inhibition of AtMSH6. For this purpose the full-length cDNA encoding AtMSH3 is isolated from plasmid pPF66 by digestion with SmaI and is cloned into the SmaI site of the DMC1 expression cassettes described in A(i).

Z.

...

E. Selection of Recombination markers on homeologous chromosomes of Arabidopsis thaliana subspecies Landsberg erecta (Ler), Columbia (Col) and C24, respectively

E(i). Visual recombination markers in Arabidopsis th. subspecies C24:

Agrobacterium mediated transformation with a T-DNA containing a 35S-GUS gene, inactivated by insertion of a 35S-Ac transposable element (Finnegan et al., 1993, Plant Mol. Biol. 22, 625-633), had yielded a C24 line in which the T-DNA construct was integrated into chromosome 2. Genetic and molecular analysis of this line shows that the Ac transposon had excised from its T-DNA locus thereby restoring GUS activity, but had re-inserted into the chromosome at a distance of 16.4 cM, where it stayed fixed (due to disablement of Ac) within the chlorina gene. Insertional inactivation of the chlorina gene caused a bleached phenotype in those plants that were homozygous for this mutation. Because of the two linked phenotypic markers, chlorina3A:Ac and GUS, this C24 line was used in crosses to wildtype Ler for analysis of meiotic homeologous recombination on chromosome 2 in conjunction with molecular recombination markers.

35 E(ii). Visual recombination markers in Arabidopsis th. Ler:

The Ler line NW1 (obtained from NASC, Nottingham, UK) contains one recessive visual marker per chromosome. i.e. an-1 on Chr.1, py-1 on Chr.2, gl1-1 on Chr.3, cer2-1

on Chr.4. and ms1-1 on Chr.5. This line is used in crosses to wildtype C24 which expresses an MMR altering gene for analysis of meiotic homeologous recombination on chromosomes 1-5 in conjunction with molecular recombination markers listed in Table 1.

Other Ler lines from NASC have several visual markers in close proximity to each other on the same chromosome. When these lines are used for hybrid production, analysis of homeologous meiotic recombination may make use entirely of visual recombination markers. Particularly suitable for crossing to C24 wildtype that is expressing a MMR altering gene are the following Ler lines:

NW22: relative markers are dis1 - (4 cM) - ga4 - (11 cM) - th1 on chromosome 1.

NW10: relevant markers are tz-201 - (5 cM) - cer3 on chromosome 5.

NW134, relevant markers are ttg - (4 cM) - ga3 on chromosome 5.

NW24 (abi3-1) and NW64 (gl1-1). When present in the same plant on chromosome 3. abi3-1 and gl1-1 are 13 cM apart. Since this marker combination is not available from NASC, we have combined these markers by crossing line NW24 to line NW64. The F1 offspring were allowed to self-fertilise and to produce F2 seeds. F2 Plants which carry both markers as homozygous traits on the same chromosome can be identified firstly, by germinating F2 seeds on germination medium containing selective concentrations of abscisic acid, and subsequently, by identifying among the abscisic acid resistant plants those individuals which show the glabra phenotype.

20 E(iii) Molecular recombination markers in Col, Ler and C24:

The genome of Arabidopsis thaliana is interspersed with unique base sequences arranged as simple tandem repeats. Allelic repeats can vary in length between different Arabidopsis subspecies and when amplified by PCR yield diagnostic DNA products of different length named Simple Sequence Length Polymorphisms (SSLPs). Many SSLPs have been genetically mapped and have been assigned to unique chromosome locations on the recombinant inbred map (Bell and Ecker, 1994, Genomics 19, 137-144; Lister and Deans lines, Weeds World 4i, May 1997).

In Table 1 are listed 28 mapped and established SSLPs between Ler and Col. A number of PCR primer pairs are described herein (SEQ ID NO:42 to SEQ ID NO:97) which also yielded SSLPs between C24 and Ler (19 SSLPs) or between C24 and Col (25 SSLPs), respectively. Polymorphic SSLPs can be used as molecular markers in the analysis of homeologous recombination between genomes from these subspecies.

The PCR reactions (25 μL) were carried out over 35 cycles (15 seconds at 94°C, 30 seconds at 55°C and 30 seconds at 72°C), with 0.25 U Taq DNA polymerase and 0.6 μg genomic DNA in reaction buffer containing 2 mM MgCl₂. PCR products were separated by agarose gel electrophoresis (4% ultra high resolution agarose) and visualised by ethidiumbromide staining. The results from the PCR experiments are summarised in

Table 1, which also shows the sequence of PCR primers, primer annealing temperature (Tm), PCR product length and chromosome location of SSLP (with respect to the RI map of May 1997, Weeds World 4i).

F. Production of hybrid plants

5 C24 plants heterozygous for chlorina3A:Ac/GUS are crossed as male to emasculated wildtype Ler to produce Ler/C24(chlorina3A, GUS) hybrid seeds.

Due to the heterozygosity of the C24 parent, only 50 % of hybrid plants have inherited the chlorina3A:Ac/GUS locus. The remaining 50% of hybrid plants are wildtype with respect to chlorina3A:Ac/GUS. Since the mutant locus is linked to a kanamycin resistance gene (contained on the same T-DNA as GUS) mutant plants can be pre-selected by germinating hybrid seeds on germination medium containing 50 mg/L kanamycin.

Ler plants homozygous for the five chromosome markers are male sterile (ms1-1) and are crossed without emasculation to wildtype C24 to produce Ler(an-1, py-1, gl1-1, cer2-1, ms1-1)/C24 hybrid seeds.

Other Ler plants, which are male fertile, are crossed after emasculation of the female parent to produce Ler/C24 hybrid seeds.

G. Introduction of MSH3 and MSH6/3 antisense genes into Arabidopsis and analysis of meiotic homeologous recombination

(i) Transformation of hybrid plants and analysis of homeologous meiotic recombination

The plant transformation vectors comprising the antisense genes described in (B) and (C) above are introduced into Agrobacterium tumefaciens strain AGL1 (Lazo et al., 1991, Bio/Technology 9, 963-967) by electroporation. Recombinant Agrobacterium clones are selected on LB medium containing 50 mg/L rifampicin and 100 mg/L carbenicillin. Selected clones are used to infect roots of Arabidopsis hybrid plants (described in (F) above) using the root transformation protocol of Valvekens et al. (1988, PNAS 85, 5536-5540) except that shoot and root inducing media contain hygromycin (10 mg/L) instead of kanamycin.

Plants regenerated from roots of hybrid plants are genetic clones of root donating plants and therefore are again genetic hybrids of two Arabidopsis subspecies described in 30 (F). However, in contrast to the root donating plants, the regenerated hybrid plants also contain the introduced transgene and the co-introduced hygromycin resistance gene with the latter allowing these plants to grow on hygromycin containing culture medium.

Hygromycin resistant plants are then allowed to enter the reproductive phase and to produce gametes by meiotic divisions of microspore and megaspore mothercells. At meiosis, the *DMC*1 promoter is activated and can direct the expression of antisense genes described in (B) and (C) above, leading to decreased *MSH*6 and/or *MSH*3 gene

expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3 protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between homeologous chromosomes.

Transgenic plants are then allowed to self-fertilise and to produce seeds. These 5 seeds (F2 seeds with respect to hybrid production), and the plants derived therefrom, carry the homeologous recombination events which can be identified by using the visual and molecular recombination markers described in (E) above.

In case of homeologous recombination between chromosomes of Ler and C24(chlorina3A:Ac, GUS), the analysis concentrates on chromosome 2 by selecting plants showing the visual phenotypic marker chlorina. This marker thus serves as a reference point as it indicates that respective chromosomes 2 originate from C24. Other markers, such as GUS or molecular markers, on chromosome 2 may then be used to identify chromosomal regions which are derived from the Ler chromosome as a result of homeologous recombination. F2 plants of control transformants not expressing the antisense gene(s) can be analysed in parallel and the results can be used for comparison to homeologous recombination results obtained in antisense plants.

(ii) <u>Transformation of C24 wildtype</u>, hybrid plant production and analysis of homeologous meiotic recombination

Introduction of MMR altering genes into wildtype C24 is done using the root transformation protocol as described in G(i) for transformation of hybrid plants. Transformed plants are selected by resistance to either 10 mg/L hygromycin (in case of transformation with T-DNA's derived from pNOS-Hyg-SCV) or to 50 mg/L kanamycin (in case of transformation with T-DNA's derived from pBIN19).

Transgenic plants are then allowed to self-fertilise and to produce seeds (T1 seeds).

Segregation of the antibiotic resistance gene in the T1 population then indicates the number of transgene loci. Lines with a single transgene locus (indicated by a 3:1 ratio of resistant:sensitive plants) are selected and are bred to homozygosity. This is done by collecting selfed seeds (T2) from T1 plants and subsequent testing of at least four independent T2 seed populations for segregation of the antibiotic resistance gene. T2 populations which do not segregate the antibiotic resistance gene are assumed to be homozygous for both the resistance gene and the linked MMR altering gene.

C24 plants homozygous for the MMR altering gene are then crossed to Ler lines homozygous for recessive visual markers (see E(ii)) to produce C24/Ler hybrid plants as described in (F). These F1 hybrids are then allowed to enter the reproductive phase and to produce gametes by meiotic division of microspore and megaspore mothercells. At meiosis, the DMC 1 promoter is activated and can direct the expression of antisense or sense genes described in (B), (C) and (D) above, leading to decreased MSH6 and/or MSH3 gene expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3

protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between the homeologous chromosomes of C24 and Ler. Recombination events are then scored in the F2 generation.

For recombination analysis, the hybrid plants are allowed to self-fertilise and to produce F2 seeds. F2 plants are then preselected for a first visual marker. Since this marker is recessive, its visual presence indicates homozygosity for Ler DNA at the relevant locus. Those F2 plants which show this first visual marker are then analysed for the presence or absence of a second visual marker which in the Ler parent is closely linked to the first marker. Absence of the second visual marker indicates recombination between the relevant C24 and Ler chromosomes between the first and second marker. The frequency of recombination in transgenic hybrids is compared to the recombination frequency in control hybrids not expressing the MMR altering gene.

Examples of recombination analysis are the following.

The Ler line NW22(dis1. ga4. th1) is used for crosses to transformedC24.

F2 plants are preselected first for thiamine requirement (th1) and then are further analysed for re-appearance of wildtype height (loss of ga4) and/or re-appearance of normal trichomes (loss of dis1) as a result of recombination.

The Ler line NW10(1z-201, cer3) is used for crosses to transformed C24.

F2 plants are then preselected first for thiazole requirement (tz) and then are further analysed for re-appearance of normal, i.e. non-shiny stems (loss of cer3) as a result of recombination.

The Ler line NW134 (ttg, ga3) is used for crosses to transformed C24. F2 plants are first preselected for dwarfish appearance (ga3) and are then analysed for re-appearance of trichomes (loss of ttg) as a result of recombination.

Ler plants homozygous for abi3-1 and gl1-1 are used for crosses to transformed C24. F2 plants are first preselected for their ability to germinate in the presence of abscisic acid and are then analysed for re-appearance of trichomes on the leaves (loss of gl1-1) as a result of recombination.

In the case of homeologous recombination between transformed C24 and the Ler line NW1 (an-1, py-1, gll-1, cer2-1, msl-1), recombination analysis is similar the one described above, except that the second marker is not a visual marker but has to be a molecular marker. This is because the Ler parent carries only one visual marker per chromosome.

		TABLE 1: S	ABLE 1: SSLP Markers in Arabidopsis thaliana Subspecies	haliana S	Subspecies		
Chromosome	RI Map Position	PCR Primer Pair	Primer Sequence	Tm l°Cl	length/COL	length/LER	length/C24
_	2.3	ATEATI F	GCCACTGCGTGAATGATATG	57.8	172	162	162
		ATEAT! R	CGAACAGCCAACATTAATTCCC	58.2			
·							
_	9.3	NGA63 F	AACCAAGGCACAGAAGCG	57.3	=	89	120
		NGA63 R	ACCCAAGTGATCGCCACC	59.6			
-	40.1	NGA248 F	TACCGAACCAAAAGG	56.1	143	129	no amplific.
		NGA248 R	TCTGTATCTCGGTGAATTCTCC	58.2			
	81.2	NGA128 F	GGTCTGTTGATGTCGTAAGTCG	60.1	180	8	no amplific.
		NGA128 R	ATCTTGAAACCTTTAGGGAGGG	58.2			
	81.2	NGA280 F	CTGATCTCACGGACAATAGTGC	1.09	105	85	85
		NGA280 R	GGCTCCATAAAAGTGCACC	87.8			
	111.4	NGAIII F	CTCCAGTTGGAAGCTAAAGGG	09	128	791	170
		NGAIII R	TGTTTTTAGGACAAATGGCG	70			
=	ca. 7.5	NGA 168 F	CCTTCACATCCAAAACCCAC	57.8	213	217	208
		NGA 168 R	GCACATACCCACAACCAGAA	87.8			

	ca 48	NGA1126L	CGCTACCCTTTTCGGTAAAG	57.8	161	199	961
=		NGA1126R	GCACAGTCCAAGTCACAACC	6.65			
						_	
=	62.2	NGA361L	AAAGAGATGAGAATTTGGAC	51.7	114	120	114
		NGA361R	ACATATCAATATATTAAAGTAGC	49.5			
11	73	NGA168 F	TCGTCTACTGCACTGCCG	59.6	151	135	135
		NGA168 R	GAGGACATGTATAGGAGCCTCG	61.9			
=	ca. 77	AthBIO2 L	TGACCTCCTCTTCCATGGAG	59.9	141	209	139
		AthBIO2 R	TTAACAGAAACCCAAAGCTTTC	54.5			
=	ca. 83	AthUBIQUE L	AGGCAAATGTCCATTTCATTG	54.1	146	148	148
		AthUBIQUE R	ACGACATGCCAGATTTCTCC	57.8			
100	3.4	NGA172 F	AGCTGCTTCCTTATAGCGTCC	09	162	136	140
		NGA172 R	CATCCGAATGCCATTGTTC	55.4			
=	12.8	NGA126 F	GAAAAACGCTACTTTCGTGG	56.1	611	147	no amplific.
		NGA126 R	CAAGAGCAATATCAAGAGCAGC	58.2			
	17.5	NGA162 F	CATGCAATTTGCATCTGAGG	55.8	107	89	no amplitic.
		NGA162 R	CTCTGTCACTCTTTTCCTCTGG	1.09			
				•			

=	8.18	NGA6 F	TGGAITTCTTCCTCTTCAC	56.1	143	123	143
		NGA6 R	ATGGAGAGCTTACACTGATC	1.95			
١٨	19.8	NGA12 F	AATGTTGTCCTCCCTCCTC	59.9	247	234	220
		NGA12 R	TGATGCTCTCTGAAACAAGAGC	58.2			
			•				
١٨	24.1	NGA8 F	GAGGCAAATCTTTATTTCGG	56.1	154	861	061
		NGA8 R	TGGCTTTCGTTTATAAACATCC	54.5			
	٠						
١٧	102	NGA1107 L	GCGAAAAACAAAAAATCCA	50.2	150	140	140
		NGA1107 R	CGACGAATCGACAGAATTAGG	88			
·							
^	8.11	NGA225 F	GAAATCCAAATCCCAGAGGG	58	611	189	611
		NGA225 R	TCTCCCCACTAGTTTTGTGTCC	1.09			
^	16.7	NGA249 F	TACCGTCAATTTCATCGCC	55.4	125	115	115
		NGA249 R	GGATCCCTAACTGTAAAATCCC	58.2			
							-
^	6.61	CA72 F	AATCCCAGTAACCAAACACACA	56.3	124	110	110
		CA72 R	CCCAGTCTAACCACGACCAC	6.19			
,>	20	NGA151 F	GTTTTGGGAAG1TTTGCTGG	55.8	150	120	130
		NGA151 R	CAGTCTAAAAGCGAGAGTATGATG	58.6			

				1 07	157	123	0.1
>	24	NGA 106 F	GITAIGGAGIIICIAGGCACG	1.00	//:-	=-	
		NGA 106 R	TGCCCCATTTTGTTCTTCTC	55.8			
3	37.8	NGA139 F	AGAGCTACCAGATCCGATGG	59.9	174	132	132
A	0.10	NGA139 R	GGTTTCGTTTCACTATCCAGG	55.8			
							000
>	80	NGA76 F	GGAGAAATGTCACTCTCCACC	£.1	231	> 250	300
		NGA76 R	AGGCATGGGAGACATTTACG	57.8			
		1 101 03112.	CTCCACCAATCAAATC	55 K	148	156	146
>_	1.10	AIRSOIN E	בוריביים וביים				
		ATHSO191 R	TGATGTTGATGGAGATGGTCA	53.7			
							·
	21.7	NGA129 F	TCAGGAGGAACTAAAGTGAGGG	1.09	7.11	179	172
• •		A 00 A 20	CACACTGAAGATGGTCTTGAGG	1.09			
		110011 IV					

CLAIMS

- 1. An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.
- 5 2. A DNA molecule according to claim 1 wherein said polypeptide is homologous to a mismatch repair polypeptide of a yeast or of a human.
 - 3. A DNA molecule according to claim 1 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 4. An isolated and purified polypeptide functionally involved in the DNA no mismatch repair system of a plant.
 - 5. A polypeptide according to claim 4 which is homologous to a mismatch repair polypeptide of a yeast or of a human.
- 6. An isolated and purified polypeptide selected from the group consisting of a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18), a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30), polypeptides homologous to a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18) and polypeptides homologous to a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30).
- 7. An isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.
- 8. A DNA molecule according to claim 7 comprising a polynucleotide sequence encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence.
- 9. A DNA molecule according to claim 8 wherein said polynucleotide is capable 30 of interfering with the expression of a plant polynucleotide sequence is a sense polynucleotide, an antisense polynucleotide or a ribozyme.
 - 10. A DNA molecule according to claim 7 comprising a polynucleotide sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

- 11. A DNA molecule according to claim 10 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 12. A DNA molecule according to claim 10 further comprising a regulation element capable of causing overexpression of said polypeptide in a cell of said plant.
 - 13. A chimeric gene comprising:

5

a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant; and

at least one regulation element capable of functioning in a plant cell.

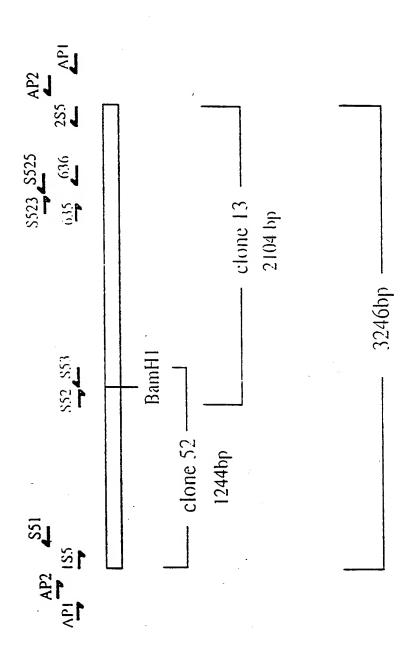
- 14. A chimeric gene according to claim 13 wherein said regulation element is selected from constitutive, inducible, tissue type specific and cell type specific promoters.
- 15. A chimeric gene according to claim 13 comprising a DNA sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant, wherein said regulation element is capable of causing overexpression of said polypeptide in a cell of said plant.
- 16. A chimeric gene according to claim 13 wherein said regulation element is selected from the group consisting of 35S, NOS, PR1a, AoPR1 and DMC1.
 - 17. A plasmid or vector comprising a chimeric gene according to any one of claims 13-16.
 - 18. A plant cell stably transformed, transfected or electroporated with a plasmid or vector according to claim 17.
- 25 19. A plant comprising a cell according to claim 18.
 - 20. A plant according to claim 19 selected from plants of the families Brassicaceae, Poaceae, Solanaceae, Asteraceae, Malvaceae, Fabaceae, Linaceae, Canabinaceae, Dauaceae and Cucurbitaceae.
- 21. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a DNA molecule according to any one of claims 1-3 or 7-12 and causing said DNA sequence to express said polynucleotide or said polypeptide.
 - 22. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a chimeric gene

according to any one of claims 13-16 and causing said DNA sequence to express said polynucleotide or said polypeptide.

- 23. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a plasmid or vector according to claim 17 and causing said DNA sequence to express said polynucleotide or said polypeptide.
- 24. A process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred.
- 25. A process according to claim 24 wherein a first gene is incapacitated in said first plant, a second gene is incapacitated in said second plant, and said first and second genes are incapacitated in said hybrid plant thereby altering the mismatch repair system of said hybrid plant.
 - 25. A process according to claim 25 wherein said incapacitation of the mismatch repair system of said hybrid plant is reversible.
- 26. A process according to claim 24 wherein a new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait is observable in at least one of said offspring plants.
- 27. A process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.
- 28. A process according to claim 27 wherein said step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene according to claim 13 and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant, cell or cells.
 - 29. A process according to claim 28 comprising inactivating an MSH3 gene and/or an MSH6 gene of said plant.
- 30. A process according to claim 28 comprising inactivating an MSH3 gene and an MSH6 gene of said plant.

- 31. A process according to claim 27 comprising at least partially inactivating the mismatch repair system of said plant in a predetermined cell type or in a predetermined tissue of said plant.
- 32. A process according to claim 31 further comprising restoring mismatch repair 5 in said cell type or said tissue.
 - 33. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule according to claim 1 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
- 34. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO: 18 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
 - 35. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.

Figure 1



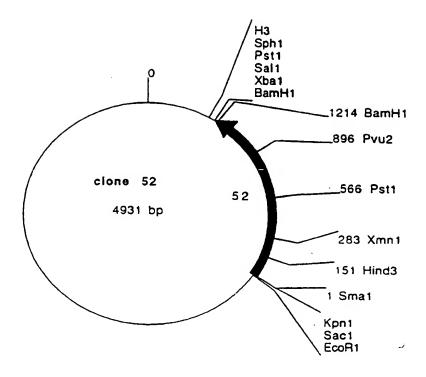


Figure 2

Comments/References: 52= 3' side of S5 (AtMSH3) 1244bp in pUC18/Sma1

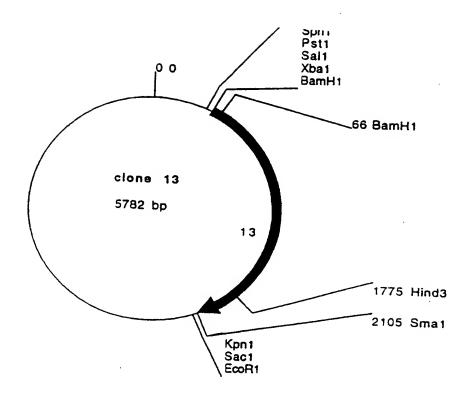


Figure 3

C mments/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Sma1

	٥	ľ
	١	
	2	į
	E	3
	ì	4
1		

9 80	144	204 35	264 55	324 75	384 95	444	504 135	564 155	624 175	684 195	744	804 235	864 255	924 275
3AGG	CCC	CCG P	CTC L	GAT D	GTT V	AAG K	GGA G	TTC F	GCA A	GCA A	GAA E	AAT N	GAA E	GTT V
MCG	GCT A	CCA P	CAC H) -	TAT Y	CTA L	TTC	ZAT N	AAT N	GGT G	CTT L	AGT S	ATT I	GTT V
TTA	TTC F	CCG	GAC	GTA V	GAA E	GAG E	TTC	CAC H	GTG V	CAT H	ACG T	CAG O	GGT G	gaa E
TCT	TTC F	ACA T	TCC S	CCA P	GAG E	GTG V	AGA R	GAT D	CTG L	TCC S	9 9	TCA S	TGT C	GGT
CAT	CGT R	TCA S	CTC L	AAC N	000 P	GTG V	TAC Y	ATG M	AGA R	AAG K	X AAA	GGT G	၁၁၁	ACA T
agcegegaaaattggcaacccaagttegecatagccacgaccacgaccttccatttctcttaaacggagga	TCT	TCA S	CTT L	CAA O	TCG S	CAA	AGG R	CAT H	AGA R	ATT I	ACC T	TTT F	TTA L	TCG S
ACGAC	ATT I	GAA E	AAG K	ACT	CCC P	CAG Q	TAC	GC'F A	GTG V	gcc A	TAT Y	GGT G	ACA T	ATT I
SACCI	ACG T	gcc A	CGT R	CAC H	GAA E	G GZ	66T 6	TAC Y	CAT H	GCA A	TTG L	GAA E	GAG E	GAA E
CACC	CAG	GT'A V	AAG K	CCT P	CTG L	TTG L	GTT V	AT'F I	TTC F	ACT T	GCG A	GAA E	TCG S	GTT V
ATAGO	CAG Q	ეეე გ	TCC S	TCT S	TTT	CCA P	GAA E	GGT G	AAT N	GAA E	TCG S	GGT G	AAG K	ე ე
, 2000	AAG K	AAT N	CCT	CTT L	AGA R	ACA T	GTG V	TTG L	TTG L	ACT	CTG I.	GGT G	GrT V	GTT V
GTTC	CAA	CCG	TCT S	AAA K	CAG 0	TAC Y	ATG M	GTG V	CGA R	CAG	GCA G	TGT	AGA R	GTT V
CCA	A.A.G.	GAA E	TTC F	CCT P	CTC L	A.A.A.A	TTG L	CGC R	TTT F	AAG K	CGG R	GGT G	GAG E	GGT G
CAAC	ງ ວງງ	CAC H	TCC S	AAG K	TTT E	AGG R	GTT V	GCA A	ACA T	GTG V	TTC F	GGT G	GAT D	GTC V
ATTG	ATG M	ACT T	GTA V	A AAA	AGA R	TCG S	GTG V	GCA A	CCA P	GTA V	TTT F	AGT S	GTG V	AGA R
SAAA	MTT	000 P	ACT	000 P	a G	TCA	GAT D	ATC I	GTG V	GGT G	CCT P	ATA I	GTT V	GTT V
9090	ATAAAGCAATT	TCC S	200 V	TCA S	CAC H	TCA S	CCA P	GAG E	AGT S	ATT I	၁၅၅	GAT D	TGT C	GAT D
AGCG	ATA	X A	TCC S	6CG	TTA L	ACG T	TAC	928 4	GCG	AAG K	ACC T	GAG E	GTT V	TTT F
CCTAAGAA	ACGA	CCC	ATA I) 4	AAT	GAA	AAG K	GAC	ACG	TAC Y	CGG R	GCT A	TTG L	AGT S
CCTA	GATTACG	A A	AAG K	200 4	CCC	CCC	AGC S	GAA E	ATG M	GGA G	AAC N	GCG A	T'IC F	ATG M
-	81 1	145 16	205	265 56	325 76	· 385 96	445 116	505 136	565 156	625 176	665 196	745	805	865 256

925 276	TAT	GAA E	GAG E	TTC F	AAT N	GAT D	AAT N	TTC F	ATG M	AGA R	AGT S	GGA G	TTA L	GAG E	GCT A	GTG V	ATT I	TTG L	AGC S	TTG L	984 295
985 296	TCA	CCA P	GCT	GAG E	CTG L	TTG L	CTT L	၁၅၅	CAG Q	CCT P	CTT L	TCA S	CAA	CAA	ACT	GAG E	AAG K	TTT F	TTG L	GTG V	1044 315
1045 316	GCA A	CAT M	GCT A	G G G	CCT P	ACC T	TCA S	AAC N	GTT V	CGA R	GTG	GAA E	CGT R	GCC A	TCA S	CTG L	GAT D	TGT	TTC	AGC S	1104
1105 336	AAT N	GGT G	AAT	GCA	GTA V	GAT	GAG	GTT V	ATT I	TCA S	TTA L	TGT C	GAA E	AAA K	A'rc I	AGC S	GCA A	GGT G	AAC	TTA L	1164 355
1165 356	GAA	GAT D	GAT D	A A	GAA	ATG	AAG K	CTG	GAG E	GCT A	GCT A	GAA E	A AAA	GGA	ATG M	TCT	TGC	TTG L	ACA	GTT V	1224 375
1225 376	CAT H	ACA	ATT	ATG M	AAC N	ATG M	CCA P	CAT H	CTG L	ACT T	GTT V	0 0 S	GCC A	CTC L	GCC A	CTA	ACG T	TTT F	1GC	CAT H	1284 395
1285 396	CTC L	A A	CAG O	TTT F	GGA G	TTT F	GAA E	AGG R	ATC I	CTT	TAC Y	CAA	ეეე ეეე	9 8 8	TCA S	TTT F	CGC R	TCT	TTG L	TCA S	1344 415
1345 416	AGT	AAC N	ACA	GAG E	ATG M	ACT T	CTC	TCA	GCC A	AAT	ACT T		CAS	CAG Q	TTG L	GAG	GTT	GTG V	X AA	AAT N	1404 435
1405 436	AAT	TCA	GAT D	GGA G	TCG S	GAA E	TCT S	၁ဗ္ဗာ	TCC	TTA	TTC F	CAT H	AAT	ATG M	AAT	CAC	ACA	CTT	ACA	GTA V	1464 455
1465 456	TAT Y	GCT	TCC S	AGG R	CTT L	CTT L	AGA R	CAC H	TGG W	GTG V	AC'F T	CAT H	CCT	CTA	7.GC C	GAT	AGA .	AAT	TTG	ATA I	1524 475
1525 476	TCT	GCT A	CGG R	CTT L	GAT D	GCT A	GTT V	TCT	GAG E	ATT I	TCT S	GCT A	TGC	ATG	GGA G	TCT	CAT	AGT	TCT	TCC	1584 495
1585 496	CAG Q	CTC L	AGC S	AGT S	GAG E	TTG L	GTT V	GAA E	GAA	GGT G	TCT	GAG	AGA R	GCA A	ATT	GTA '	TCA	CCT (GAG	TTT F	1644 515
1645 516	TAT Y	CTC	GTG V	CTC L	TCC S	TCA S	GTC V	TTG L	ACA T	GCT	ATG M	TCT	AGA R	TCA	TCT	GAT	ATT (I	CAA	CGT (GGA G	1704 535
1705 536	ATA I	ACA	AGA R	ATC	TTT F	CAT H	000 8	ACT	GCT A	A AA	GCC A	ACA	GAG	TTC	ATT	GCA (GTT 1	ATG (GAA (GCT A	1764 555
1765 556	ATT	TTA L	CTT L	900 4	999	AAG K	CAA O	ATT	CAG O	ccc crr R L		GGC ATA AAG CAA G I K Q Q 4 (Continued)	ATA I onti	AAG K •	cya o	GAC D	TCT (GAA 1	ATG .	AGG R	1824 575

ATG C	CAA	TCT	GCA A	ACT O	GTG (CGA 1 R S	TCT S	ACT (CTT 1	TTG A	AGA A	AAA 1 K	TTG A	ATT T I S	TCT G	GTT A	ATT	TCA S	1884 595
GTT GTG GTT GAC	GTT GAC	GAC			AAT O	GCC (GGA G	AAA (CTT (CTC 1	TCT (GCC (A	CTA 1	AAT A	AAG G K	GAA G	929 P	GCT A	1944 615
GGT GAC TTG CTC G D L L	TTG CTC L L	CTC		_ U ⊃	GAC 1	ATA (I	CTA	A'I'C I	ACT T	TCC /	AGC (S	GAC (CAA	TTT C	CCT G	GAG C	CTT (GCT A	2004
GCT CGC CAA GCA GTT T' A R Q A V L	CAA GCA GTT Q A V	GTT V		<u> </u>	TTA (GTC /	ATC I	AGG (GAA 7	AAG (CTG (GAT 7	TCC 3	TCG A	ATA G	GCT 1	TCA	TTT E	2064 655
AAG CTC GCT ATT K L A I	GCT ATT A I	ATT 1		ഠ∝	CGA I	AAT N	TTG	GAA ,	TTT	CTT (CAA	GTG '	TCG (666 #	ATC A	ACA C	CAT	TTG L	2124 675
CTG CCC GTT GAT L P V D	CCC GTT GAT	GAT D		ည်	TCC	AAG (G1C V	CCT	ATG	N N	TGG (GTG /	AAA K	GTA P	AAT A	AGC I	ACC	AAG K	2184 695
ACT ATT CGA TAT CAT CCC T I R Y H P	CGA TAT CAT R Y H	CAT H		ပ္ပ		CCA (GAA E	A'FA I	GTA V	CC1.	 	115	GAT O	GAG C	CTA (GCT (CTA	GCA A	2244 715
GAA CAT CTT GCC ATT GTG E H L A I V	CTT GCC ATT L A I	ATT I		GT V		AAC (CGA R	GCT	TCG	TGG W	GAT	AGT S	TTC . F	CTC A	AAG 1 K	AGT S	TTC	AGT S	2304 735
TAC ACA GAT TTT AAG	ACA GAT TTT T D F	TTT		ξ×		GC'F	OCC A	GTT V	CAR	GCT A	CTT	GCT A	GCA	CTG (GAC 3	TGT	TTG	CAC H	2364 755
CTT TCA ACT CTA TCT AGA	ACT CTA TCT T L S	TCT S		7		AAC	AAG K	AAC	TAT Y	GTC V	CGT.	CCC	GAG . E	TTT (F	GTG (GAT (GAC	TGT C	2424 775
A GTT GAG ATA AAC ATA V E I N I	GAG ATA AAC E I N	AAC		AT I		CAG	TCT S	GGT	CGT R	CAT H	CCT	GTA V	CTG	GAG A	ACT /	ATA I	TTA	CAA	2484 795
AAC TTC GTC CCA AAT GAC N F V P N D	GTC CCA AAT V P N	AAT N		წ <u>ი</u>		ACA	ATT I	TTG L	CAT H	GCA A	GAA	999	GAA E	TAT 1 Y	7GC (CAA	ATT	ATC I	2544 815
GGA CCT AAC ATG GGA GGA G P N M G G	AAC ATG GGA N M G	GGA G		ပ္ပ် ပ		A.A.G.	AGC S	TGC	TAT Y	ATC	CGT.	CAR	GTT V	GCT 7	TTA I	ATT '	TCC	ATA I	2604 835
GCT CAG GTT GGT TCC TTT A Q V G S F	GTT GGT TCC V G S	TCC S		7	Ę.	GTA V	CCA P	9 900	TCA	TTC F	000 A	AAG K	CTG L	CAC (GTG (CTT (GAT D	GGT	2664 855
TTC ACT CGG ATG GGT GCT F T R M G A	CGG ATG GGT R M G	ATG GGT M G		₽ C.	⊢	TCA	GAC D	AGT ATS S I	ATC I	CAG 0	CAT H onti	GGC AGP Ge R inued)	AGA R d)	S.	ACC T	F	CTA L	GAA E	2724 875

2725 876	GAA	TTA L	AGT	GAA E	GCG A	TCA S	CAC H	ATA I	ATC I	AGA R	ACC	TGT	TCT S	TCT (S	CGT 3	TCG (CTT	GTT	ATA I	TTA L	2784 895
785 896	GAT D	GAG E	CTT L	GGA G	AGA R	၁၅၅	ACT T	AGC S	ACA T	CAC H	GAC	GGT	GTA (gcc A	ATT (GCC A	TAT	GCA	ACA T	TTA L	2844 915
845 916	CAG Q	CAT H	CTC L	CTA	GCA	E AA	A.A.G K	AGA R	TGT	TTG L	GTT V	CTT	TTT (GTC V	ACG (CAT	TAC	CCT P	GAA E	ATA I	2904 935
2905 936	GCT A	GAG E		AGT	AAC N	GGA G	TTC F	CCA P	GGT	TCT S	GTT	999	ACA T	TAC	CAT (GTC V	TCG	TAT Y	CTG L	ACA T	2964 955
965 956	TTG L	CAG 0	AAG K	GAT D	A A	့ ၁၅၅ ၂	AGT S	TAT Y	GAT D	CAT	GAT D	GAT	GTG .	ACC	TAC	CTA	TAT	AAG K	CTT	GTG V	3024 975
025 976	CGT R	GGT	CTT L	1 60	AGC S	AGG R	AGC S	TTT F	GGT G	TTT F	A.A.G. K	GTT	GCT	CAG	CTT	gcc A	CAG	ATA I	CCT	CCA P	3084 995
085 996	TCA S	TGT C	ATA I	CGT R	CGA R	9	ATT	TCA S	ATG M	GCT A	GCA A	AAA K	TTG	GAA	GCT A	GAG	GTA V	CGT R	GCA A	AGA R	3144
145 016	GAG E	AGA R	AAT	ACA T	င် က	ATG	GGA	GAA E	CCA P	GAA E	GGA G	CAT	GAA	GAA	CCG P	AGA R	၁၅၅	GCA	GAA E	GAA E	3204 1035
205 036	TCT	ATT I	TCG S	GCT A	CTA L	GGT	GAC	TTG L	TTT F	GCA A	GAC	CTG L	× ×	T'IT F	GCT	CTC L	TCT	GAA	GAG E	GAC	3264 1055
265 056	CCT P	TGG ¥	¥ ×	GCA A	TTC F	GAG E	TTT F	TTA L	AAG K	CAT H	GCT	TGG W	AAG	ATT 1	GCT	299	X AA	ATC I	AGA R	CTA L	3324 1075
325 076	A A	CCA P	ACT T	TGT C	TCA	TTT F		TTTA	WTCI	TAAC	ATTA	TAGC	AACT	GCNA	GGTC	TTGA	TCAT	ctgr	TAGT	TGA TTTAATCTTAACATTATAGCAACTGCAAGGTCTTGATCATCTGTTAGTTGCG	3397 1082
398 1	TACI	TACTAACTT		ATG TC	TGT A1 C I	ATT AC I S	AGT A1 S I	ATA AC I T	ACA AG T R	AGA AAA R K		GAG AAT E N	T TA	TAG AGAG	AG M	ATG G M D	GAT T D S	TCT A	AAT C	900 900	3458 5
459 6	GTG >	TTG L	CAG O	TAC Y	ATC	TTT	TCT	CCA P	ددد اله	GCA A igus	GCA TAA A A igure 4	44	uaaaaaaaaaaa (Continued)	AAAA	AAAA d)	AAAA	AAAA	AAAA	\$		3522 16

:

igure 5

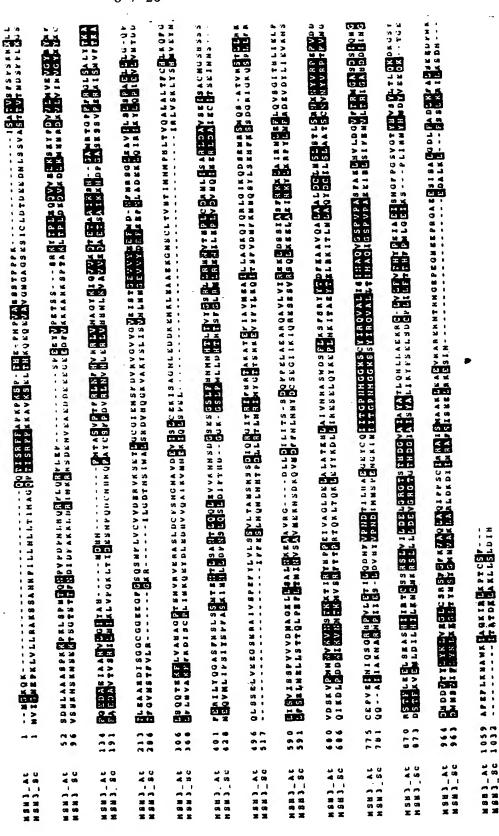
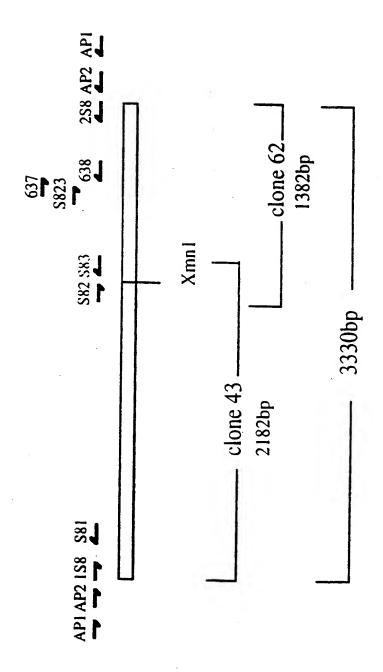


Figure 6



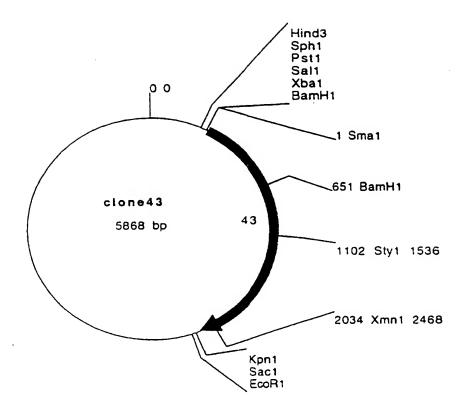


Figure 7

Comments/References: 43= 5' side of S8 (AtMSH6) 2182 bp in pUC18/Sma1

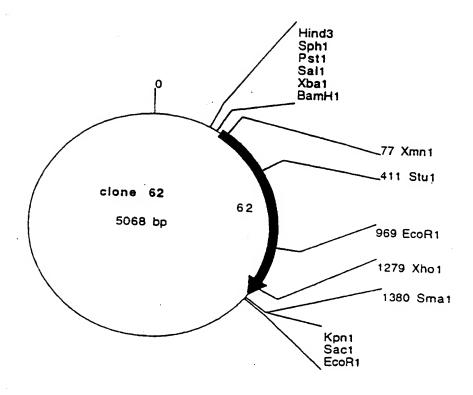


Figure 8

Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Sma1

	153 4	213 24	273	333 64	393 84	453 104	513 124	573 144	633	693 184	753 204	813 224	873 244	933 264
aaaagttgagccctgaggagtatcgtttccgccatttctacgacgcaaggcgaaatttttggcgccaatctttcccccc	_U	TCC	999	AGA R	CCG P	A AAA	GAT D	¥ ×	GAA E	GCT	GTA V	GAA E	AGA R	TTC
rrcc	c cAG	GTT V	GAA	GTT V	AAG K	GTA V	AAT N	GGT G	GTA V	CGT R	CCT P	A.A.A	AAT N	GTT
ATCT	ဂ င က က	TTG	AAG R	GAG	TTT F	TTT	CTG L	AAT N	TCA S	CCA P	GTT V	A.A.G K	9 8 8	GAT D
GCCA	G CAG	GGT	GTG V	GAT D	GGA	AAG K	CCG P	AAT N	AGA R	CGT R	AAG K	GAG E	GAT D	CCT
TGGC	TCTCTCAGCTCAAAACATCGTTTCTCTCTCTCTCTCTCACAATTCCAAAAA ATG	AAG K	AAT	GTC V	TCT	CAT H	GTT V	TCC S	CTT L	ATG M	GAT D	GGA G	AGG R	CCA
TTTT	AAAA	ACG T	TTT F	TCT	000 P	ATG M	GTT V	CGT R	GAA E	9 9 9	GAG E	TGT C	ATC I	ATA I
AAAA	TTCC	ACT T	CGA R	AAA K	CTG L	ATT I	GAT D	ن الدائيل	GCT	CCA P	AAG K	GTT V	CGA R	CAC H
ງງງງ	ACAA	GCG	CCA P	TCG S	GTC V	AAT N	GAA	CAA	AGA R	ACA	TTT F	CCG P	TCT S	TTA L
GCAA	TCTC	GCG A	GGA G	GTT V	CGT R	TCC S	CGA R	CCT P	ე ე	GAA E	ACT T	GAT D	TCT	ACC T
CGAC	rcrc	ACC T	GGA G	GCT A	cgt R	TTC F	AGC S	A'l'ſ I	AGT S	CCA P	ATG M	CAG Q	GAG E	A.A.G K
TCTA	TCAC	ردد ه	AGC S	TTT	CCG P	CTG L	AGG R	G1"r V	TTC	GGT G	GAA E	CTC L	CTT	AGA R
CATT	TCTC	A AAA	၁၅၅	CGT R	GTT V	TCC S	GAG E	GA'I' D	AGT S	CCT P	GAT D	ATG M	TGG W	GAT D
ວ໑ວວ	TTTC	CAA	ე ე	GTA V	AAG K	TCG S	GGA G	z A z	TTT F	GTT	GAG E	A AA	GAA E	TAC
GTTT	ATCG	TTC	9 999	TCT	GAG E	GCT	TCT S	GCT A	GCT A	GAT D	CTG	CTG L	TTT F	CTT L
TATC	AAAC	TTC F	ეეე ეეე	GCT	CCG P	GAT D	TGT	A AG	CAT H	၁၁၁	GTT V	AGG R	AAA K	CCC P
GGAG	CTCA	TCT	AGC S	GAC	CCA P	GST G	GAT D	ATG M	A.A.C. N	GAT D	CGA R	A AAA	ACC T	GAT D
CTGA	TCAG	TTG L	GCT	ວຽວ	ACT	225 8	CGA R	TGT C	AGA R	GTA V	AAG K	AAC	GGA G	GAT D
AGCC	TCTC	ATT I	GCT	A A A	GAT D	TCC S	GAT D	CTA L	GAA.	GGA G	TTG L	TCT S	GAA E	CCT P
STTG		TCG	GAT D		ACG T	GAA	GAT D	TCT	CAA		ეე ჯ	GAC	AAC	CGT R
AAAA	TTTCGAAT	AGA R	၁၅၅			GCT	GTC V	TCA	ACT	GAT D	TCT	TTG L	GTA V	AGA R
7	81	154 5	214 25	274	334 65	394 85	454 105	514 125	574 145	634 165	694 185	754	814	874 245

993 284	1053 304	1113 324	1173 344	1233 364	1293 384	1353 404	1413 424	1473	1533 464	1593 484	1653 504	1713 524	1773 544	1833 564
57.73	~ €	3.5	. ~ m	3.	3	4	1	~ 4	→ 4	4	5 5	53	5, 17	16
ATT I	ggt G	GGT G	GTT V	ATA I	9 999	TCA S	ATC	GAA	ACG T	GAT D	TCA	GAG E	TTT F	GAG E
GAC /	TTA L	GTT V	A AA	ACT T	ATC I	TGT C	TCC	AAG K	TAT Y	ACA T	GAA E	GGA G	ATT	CTT
ATG (GAA E	CAG Q	TAT Y	AAT N	A.A.C.	AAG K	999 999	CCA P	AAA K	GAT D	TCT	CTT	GAT	N
TAT A	GCG A	AGA R	GGA G	GCT	GGA G	CAA O	GTT V	TCT	AGG R	999	TCT	GCT	999	GTA I
GAA 1	GAT D	76C C	CGT R	GGT G	GAG E	CTA L	TGG W	GTT V	CTA	ATG M	GGT	AGT S	CAT (ATG (
AGT G	CTA L	AAA K	GCT	AGA R	AGC S	GAG E	TTT F	CAG Q	GCT A	GTA V	A A	CTT	AAG (ACG 1
AAG P	GAG E	GGA G	TTA L	GCC	GCA	ATG M	AGG R	ATG M	AAG	CAN	TTT	GCC A	CTT	CAG 1 0 ° 1
GTT A	TAT Y	GTG V	CTA L	* A	ACA T	A X	TTG L	TTG L	C A	CCA P	TAC Y	GTT V	GTA V	T GGC (G)
AGT (CTG L	GGT G	AAG K	GCA A	TCA S	ATC I	9 8	TTA L	GCA	GTA V	GGA G	GAT D	GAT	GAT D
TGG A	GAG E	AGT S	CAA	C A	CCA P	GAG E	GCT A	gcg A	GAA E	CCA P	AAC N	TGT	GAA E	A.ATT I gure S
TAT 1 Y	TAT Y	ATG M	GTG V	GAC D	ACT T	A AAA	TGT	GG.A G	AGA R	GCT	TCT S	GAA E	CTA L	AGA R Figu
CAA	TTT F	ACC T	GCA A	TCT S	TTA L	ATA I	GAC	CIT L	TCA S	TTG L	GAA E	AAT N	AAG K	CTC
AAG (AAA K	ATG M	GAG E	ACA T	GTA V	GCT A	GTT V	GCT A	CTA L	CAG Q	A'FA I	CTA L	CTA L	TGT
CAN 1	9 999	A.A.G K	GAT D	GAA E	CAG Q	CTT L	TTT F	GCT A	ეეე ე	GTA V	ATA I	GGT G	AGG R	GGT
TCA (GTG V	TGG X	ATA I	CTA L	GTT V	CTT L	GCT A	TGT C	AAA K	920 8	AAT N	GAT D	TCT	AGG R
SCA 7	AAA K	GAC	999 9	CAG	CTA L	CAT H	TTT	TCA S	AGT	ACG T	AGA R	GTT V	CTG L	TAC Y
ည် အ	TTT F	CTT	AGT	GAG E	AAG K	GTC V	GGA G	GCA A	GAC D .	TCT	GTT	GCT	CAT H	GTT V
ATG 3	TTC F	GAG E	GAA E	ATC I	AGG R	000 4	TAT Y	GAT D	TAT Y	9	SGA S		AAT	A C
X X	CTT TTC L F	A.A.G	TCT	CGA R	CCA P	GAT D	GTG V	GAT D	TTA L	G ACA (GCT A		ATT I	TAC
AAG AAA ATG K K M	GTG V	CAC	ATC	GGA	ATT C	CCT	ACT	AGC	GTG TTA TAT (TTG L	GCT A	TGG	CTA	CCA TAC P
934 265	996 285	1054 305	1114 325				1354 405	1414 425	445	1534 465	1594 485	1654 505	525	545

TAT CTT GAT AAC 1893 Y L D N 584	CTC AAA GAT 1953 L K D 604	TCA GAA AGT 2013 S E S 624		CTC GGA CGC 2073 L G R 644	CTC GGA CGC L G R AAA AAA GTG K K V	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCA GCC E A A	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCC E A A GCT GAA ACT A A	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCC E A A GCT GAA ACT A E T GTC ATT CAC V I H	CTC GGA CGC L G R AAA AAA GTG K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCC E A A GCT GAA ACT A E T GTC ATT CAC V I H TCT GCT GGA S A G	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT CTC TGT AAA L C K GAA GCC E A A GCT GAA ACT A E T GTC GGA Y I H TCT GCT GGA S A G GAA ACA O I H TCT GCT GGA S A G	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCC E A A A GCT GAA ACT A E T GTC ATT CAC V I H TCT GCT GGA S A G CAG AAA ACA Q K T GCC GAT GGT A D G	CTC GGA CGC L G R AAA AAA GTG K K V GGA ATT GAT G I D CTC TGT AAA L C K GAA GCA GCC E A A A GCT GAA ACT A E T GTC ATT CAC V I H TCT GCT GGA S A G CAG AAA ACA Q K T GCC GAT GGT A D G GCC GAT GGT A D G
CCA CTC AAA P L K		AAC TCA N S	CTG CTC L L		GGG AAA G K	GGG AAA G K AGT GGA S G	GGG AAA G K AGT GGA S G AAA CTC K L	GGG AAA G K AGT GGA S G AAA CTC K L TTC GAA	GGG AAA G K AGT GGA S G K L TTC GAA F E	GGG AAA G K AGT GGA S G AAA CTC K L TTC GAA F E E AAC GCT N A A	GGG AAA G K AGT GGA S G AAA CTC K L TTC GAA F E AAC GCT N A GGG GTC CTC TCT	GGG AAA G K AGT GGA S G AAA CTC K L TTC GAA F E AAC GCT N A CTC TCT L S AAT CAG N Q	GGG AAA G K AGT GGA S G AAA CTC K L TTC GAA F E E A A A CTC TCT L S AAT CAG N Q	GGG AAA G K AAGT GGA S G AAA CTC K L TTC GAA F E E AAC GCT CTC TCT L S AAT CAG N Q GCA GCC A A
	C TGC CAT	C.ACG GCA T.A.	'A GAA AGA E R	T CIT CTG		1.T.C.	TTC F CTT L	TTC F CTT L TCT S	TTC F CTT L TCT S GAT D	TTC F CTT L L TCT S GAT D	TTC F CTT L TCT S GAT D W	TTC CTT L L L TCT S GAT W A A A D D D D D D D D D D D D D D D D	TTC F CTT CT S S GCA A A A GCA A A A A A A A A A A A A A	TTC F. CTT CT S S GCA A A A A A A A A A A A A A A A A A A
-1 -1	T TGG ATC W I	A GAA TTC E F	A GAC T'FA D L	G CCT GCT P A		GTG AAA GG V K G	AAAA K ACT	AAA K ACT S S	AAAA K K S S S F F C C C C	AAAA K K K S S S GTG V V	AAAA K K S S S S GGTG V V ACT T T	AAAA K K K K K K K K K K K K K K K K K	AAAA K K K K K K K K K K K K K K K K K	AAAA K K K K K K K K K K K K K K K K K
S G	AGG AAT	GTT GAA	CTT CCA	GTG TTG V L	AT	: –	A TG	ATG M GAG EE	A TG M CGAG GAG CGAA	ATG AATG CAA CAA AGA AAGA A	ATC M M GAG CAA O O ACA	ATG M AGA AGA R LTT F F S S S	ATG M AGA AGA AGA AGA AGA AGA AGA AGA AGA	ATC A AGA A
	crc rra L L	GAT. GTA D V	CAC AVA	GCC TCT A S		Ö O								
0 0 0	CGA R	CTT	TAT CTC (TCA TCA (GCA TIT	نبا	F GAA E	F GAA E B AGC S	F GAA E E AGC S S	F GAA E AGC S Y Y TTT	F GAA E S S S Y Y TTT F F	E GAA E S AGC S S Y Y TTT F E E L L L TTT TTT TTT F E L L L L TTT TTT F E E E E E E E E E E E E E E E	E GAA E E S S S Y Y TTT F TTT F F C C C C C C C C C C C C C	E GAA E E S S S Y Y TTT F F C C C C C C C C C C C C C C C C
ပ	GGT AAG G K	AAA CGG K R	CAG Q	CGA R	AVA	×	C AG	K CAG Q GGA	K CAG Q GGA G CCA	K CAG Q GGA G CCA CCA	K CAG Q GGA G CCA P P GAA E	K CAG Q GGA GCA CCA P CCA CCA CAA CAA CAA CAA CAA CAA	K CAG O GGA GGAA E GGTT V AAAA K	K CAG O O O O O O O O O O O O O O O O O O O
က z	CCA ACT	ATC AAT I N	ACT GGC T G	AGC GTT S V		R V S1								· · · · · · · · · · · · · · · · · · ·
z	AGT	AGC S	ATC	TCT	CAA	a	O TTG	Q TTG L L ATA I	C L L ATA ATA I I AGC S	TTG L L I ATA AGC S S ATA I I	C TTG L L L ATA I I S S S S AGC S S S S S S S S S S S S S S S S S S S	O C C C C C C C C C C C C C C C C C C C	L L L L L L L L L L L L L L L L L L L	L L L L L L L L L L L L L L L L L L L
in the	rgr grr c v	GTA GAP V E	ATG CAA	ATC AAC	CTG AN	7	L K CTG TT	CTG TTC L L CTT CC' L P	CTG TTG CTT CC CTT CC L L P ATA GA I	CTG TTCC CTT CC' L L P L P L P ATA GA' I D CTC ACC L T L	CTG TTG CTT CC' CTT CC' L P ATA GA' I D CTC ACG L T	CTG TTG CTT CC' L P ATA GA' I D CTC AC L T L T ACC AT AGC AT S M	CTG TTG CTT CCT L L L P ATA GA' I D CTC ACC L T L T L T AGC AT S M AAA GG K G	L K CTG TTG L L CTT CCT L P ATA GAT L T CTC ACA L T ACC ATA T I ACC ATG S M AAA GGG K G CAA TTG O L
	1894 585	1954 605		2074 645	2134	999	665 2194 685	665 2194 685 2254 705	665 2194 685 2254 705 2314 725	2194 685 685 705 2314 725 2374	2194 685 685 2254 705 725 2374 745 765	2194 685 685 2254 705 725 725 745 745 765 785	2194 685 685 2254 705 2314 725 745 765 765 785 805	

2734 865	GCA A	ACA	TGT	CTG L	gcc •	GTT V	ATC	TTT F	GCC A	CAA	CTT	ე ე	TGC	TAC Y	GTG V	000 P	TGT	GAG	TCT	TGC	2793 884
2794 885	GAA E	ATC	TCC	CTC L	GTG V	GAT D	ACT T	ATC I	TTC F	ACA T	AGG R	CTT L	၁၅၅	GCA	TCT S	GAT	AGA R	ATC	ATG	ACA	2853 904
2854 905	GGA G	GAG E	AGT S	ACC	TTT F	TTG L	GTA V	GAA E	TGC C	ACT T	GAG E	ACA	GCG A	TCA S	GTT	CTT	CAG	AAT	GCA	ACT T	2913
2914 925	CAG	GAT D	TCA	CTA	GTA V	ATC I	CTT L	GAC D	GAA E	CTG L	၁၁၁	AGA R	GGA	ACT	AGT	ACT	TTC F	GAT	GGA G	TAC Y	2973 944
2974 945	သ ဗ	ATT I	GCA A	TAC Y	TCG S	GTT	TTT F	CGT R	CAC H	CTG L	GTA V	GAG E	X AA	GTT V	CAA	TGT C	CGG R	ATG M	CTC	TTT	3033 964
3034 965	GCA A	ACA T	CAT H	TAC	CAC H	CCT P	CTC L	ACC T	AAG K	GAA E	TTC F	GCG A	TCT	CAC	CCA P	CGT R	GTC V	ACC T	TCG S	AAA K	3093 984
3094 985	CAC	ATG M	GCT A	76C 0	GCA A	TTC	A AA	TCA S	AGA R	TCT	GAT D	TAT Y	o CAA	CCA P	CGT R	GGT	TGT C	GAT D	CAA	GAC D	3153 1004
3154 1005	CTA L	GTG V	TTC F	TTG L	TAC Y	CGT R	TTA L	ACC T	GAG E	GGA G	GCT A	TGT	CCT P	GAG E	AGC S	TAC Y	GGA G	CTT L	C CA	GTG V	3213 1024
3214 1025	GCA A	CTC L	ATG	GCT A	GGA	ATA I	CCA P	AAC N	CAA	GTG V	GT'F V	GAA E	ACA T	GCA A	TCA S	GGT	GCT A	GCT A	CAA	9 8	3273 1044
3274 1045	ATG	A.A.G. K	AGA R	TCA S	A'l'r I	9 999	GGA E	AAC N	TTC	AAG K	TCA S	AGT S	CAG	CTA	AGA R	TCT S	GAG	TTC F	TCA S	AGT S	3333 1064
3334 1065	CTG	CAT H	GAA	GAC D	TGG W	CTC	A.A.G	TCA S	TTG L	GTG V	GGT G	ATT I	TCT	CGA R	GTC V	GCC A	CAC	AAC N	AAT	9 4	3393 1084
3394 1085	CCC P	ATT I	၁၅၅	GAA	GAT D	GAC	TAC	GAC D	ACT	TTG L	TTT F	7GC C	TTA L	TGG W	CAT H	GAG	ATC I	A AAA	TCC	TCT	3453 1104
3454 1105	TAC	TGT C	GTT V	000 6	X A	TAA	ATG M	GCT A	ATG M	ACA T	TAA		ATCT	GAAG	CTCG	TTAA	GTCT	CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT	CCTC	TCT	3521 5
3522 1	ά Σ Ό	ATG T'	TTT A' F I	ATT CO	CCT CI	CTT A	AAA AA K	AAA TC K C	TGC TT	TTA TA L Y	TAT AT Y I	ATC AA I K	AAA AAA K K	A TTG	G TTT F	T CCT	T CGA R	ia tta L	A AAA K	4	3579 19
3580 20	AAA K	K A		AAA AAA K K	K AA	K AA	AAA K	AAA K	AAA K		Figure	<u>)</u> 6	(Continued)	penu	_						3606 28

al aman

MAN	70 PR 97 PR	
1 DEBOTO	7.0	
**> G D T G R R X T S G D T R		16-92 9202 18-92 9202 18-92 9202
NUMBER HEGODOOP STIDEN TO NOT DISTULNED CHOUSE BELLOOF THE STAND OF THE STANDED BELLOOF THE STANDED BELLOO	33	ACKEEG VOOR
RYTETOSTAVOLAPVPOVHODY DA AOVANITIS DA COSSESSION VOLHER POTALISTE ELEMENARADE VINDA ODI PPVALCELATOON DE PARTIE DE COSTAVE DE COSTA	22	ACKBEG VOOR
ENTER THE SECOND OF THE SECOND	22	1
sasq Koymay Korumbiya Lingy Ding Landa Dalah and Dalah Tus Dy Orchuy Mis Bada and and Banya Rybers Ballah and Bada Chara Chara Bada	112	10-5X:9X
Vedigvoreterstross de estrivis demetredevivies deservicos de estrogrames estrogrames de portinos de la productivi de la composition della composition della composition de la composition de la composition della	.55	1
19 PREKVERNYLPROFINTERADELFSELFENNYKKYDONDCROESKREKVOURSELFORNOMIPOPERANGOTORENKADELESSELFORNESTELFORNESTEL	::	14-929214
MARKET METERO BETTER ON THE TOTAL OF THE TOTAL ON THE TOT		ACMBES - B1

·	
TTTTTTGGTTGCTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA	50
CTATATATAAAAGAAATGAAAGATATATATTGTTTTTTCATTTATCAAAC	100
AAAACAACAAGACTTTTTTTTTTTTTTTTTTTTTTTTTT	150
GATAAACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA	200
CCTAGAACCTTCTCCATCTTCGCAAGCACAGCCTGATTAGGAACAGCTTT	250
ACCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG	300
CCAAATCCGCTTGTGACATCTTCTTCTCCAATCTCGCTTTCTGTATCATC	350
AACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC	400
TTCCAGCTTCTTCGTGTTAATCACCGGAACCGCCGTAGATTTCCCCTTTT	450
TGTTCGAACCGGCATCGAATTTCTTAACCGTTTGAACCGCGACACCGTTT	500
CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGTAGGTCTTGGCTCTTTTG	550
TTTTGATTTGTGGAGAACTACTGGTTCCCAGTCTTGTGTTACTGCTCCTG	600
GGTATCTGCTCGGCATCGTCGATGAATTGAGAGAAAGGAACAACGCGAAA	650
ATTTTATTAATCTGAGTTTTGAAATTGAGAAACGATGAAGATGAAGAATG	700
TTGTTGAGAGGATTGTGATATTTATATATACGAAGATTGGTTTCTGGAGA	750
ATTCGATCATCTTTTCTCCATTTTCGTCTCTGGAACGTTCTTAGAGATG	800
ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTGGGT	850
TGGATTCGTGGTACACCATATTATCCGATTTGGCTCAATGGTTTTATATA	900
AATTTGGTTTTCGGTTCGGTTATGAGTTATCATTAAAATTAAGCTAACCA	950
AAAATTTTCGTAAAATTTATTTCGGTTTCAATTCGGATCCCTTACTTCCA	1000
GAACCGAATTATTCSAAACCGGGGTTAGCCGAACCGAATACCAATGCCTG	1050
ATTGACTCGTTGGCTAGAAAGATCCAACGGTATACAATAATAGAACATAA	1100
ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAGTT	1150
GAGCCCTGAGGAGTATCGTTTCCGCCATTTCTACGACGCAAGGCGAAAAT	1200
TTTTGGCGCCAATCTTTCCCCCCTTTCGAATTCTCTCAGCTCAAAACATC	1250
GTTTCTCTCTCACTCTCTCACAATTCCAAAAAATGCAGCGCCAGAGAT	1300
CGATTTTGTCTTTCCCAAAAACCCACGGCGGCGACTACGAAGGGTTTG	1350
GTTTCCGGCGATGCTGCTAGCGGCGGGGGGGGGGCGGAGGCCACGATTT	1400
AATGTGAAGGAAGGGGATGCTAAAGGCGACGCTTCTGTACGTTTTGCTGT	1450
TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG	1500
TTCCGCGTCGTGTCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	1550
GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	1600
TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTCTTAATTT	1650
TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTGTATTCGTT	1700
GGAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGGAG	1750
CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	1800
ATGATGTTATTCCTCAATTTCGTTCCAATAATGGTAAAACTCAAGAAAGA	1850
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA	1900
TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC	1950
CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG	2000
GAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA	2050
GGATCCGGTTTGTGGAGAGAAGAAGAAGTAAACGAAGGAACCAAATTTG	2100
AATGGCTTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	2150
GATCCCCTTTACGATAGAAAGACCTTACACATACCACCTGATGTTTTCAA	2200

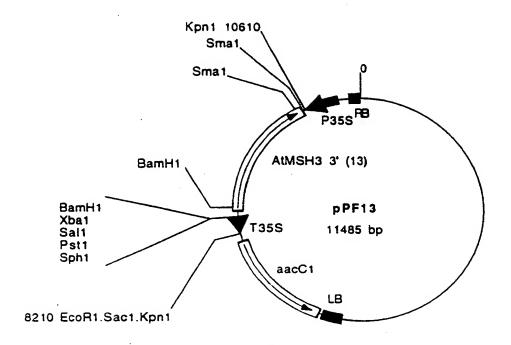
Figure 11

GAAAATGTCTGCATCACAAAAGCAATATTGGAGTGTTAAGAGTGAATATA	2250
TGGACATTGTGCTTTTCTTTAAAGTGGTTAGTAACTATTAATCTAGTGTT	2300
CAATCCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2350
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAATTAG	2400
GTCACAAGGAGCTTGACTGGAAGATGACCATGAGTGGTGTGGGAAAATGC	2450
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCTTGAATTATTGTGTCTA	2500
TAAATTTTGACACCACCTTTTGTTTCAGGTTGGTATCTCTGAAAGTGGGA	2550
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	2600
ACTTTATGGAATTCGTTTACTGCTACTTCGGCTAGGATTTAAGAAATGGA	2650
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC	2700
TTATTCGTTATATAATAAGTCTTTTCATCAAGGAGTAACAAACA	2750
GCACAATATTTGTGTGCTCACTGGCAAGGCATATATACCCAGCTAACCTT	2800
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	2850
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTGAAGCCTTGTGTTGAAA	2900
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATG	2950
CTTTCCTATTCTACTAATTATGTATTGACACACTCGTTGTTTCTTTTGAA	. 3000
AGATATAAAGTTGGACGAATCGAGCAGCTAGAAACATCTGACCAAGCAAA	3050
AGCCAGAGGTGCTAATACTGTAAGTTTTCTTGGATAGGTCAAGGAGAGTG	3100
TTGCAGACTGTTTTTGATCATTTCTTTTTTTTTGTACATTACTTTCATGCTG	-3150 -
TAATTAACTCAATGGCTATTCTGGTCTGATTATCAGATAATTCCAAGGAA	3200
GCTAGTTCAGGTATTAACTCCATCAACAGCAAGCGAGGGAAACATCGGGC	-3250
CTGATGCCGTCCATCTTCTTGCTATAAAAGAGGTTTGTTATTTACTTATT	3300
TATCTTATCATGTTCAGTTCATCCAAGTCCTGAAAAATTACACTCTTCTT	3350
TACCAATCTTCCATCAAGCTGTGTAAAGGATTTGGAATTAGAAAATCATT	3400
ATTTGATGCTTTGTTTTATATGCAAGAGGTTCCCTTGAAAAGATCTGTTT	3450
AAGATTCTTTGCACTTGAAAAATTCAATCTTTTTAAGTGAATCCCCTACT	3500
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCCT	-3550
TGTTACTGCATCCCCCTCTTTCTTAATGACCATTGTCTATGTTGTGTTTG	3600
TCTCGTGTGCTGGAGAAAATGATAGCTGATCCAAGCTGTACATTATCATG	3650
ATTAAGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG	3700
ATGTCAATTTTCCTCTGAATCTTTATTTTAGATCAAAATGGAGCTACAA	3750
AAGTGTTCAACTGTGTATGGATTTGCTTTTGTTGACTGTGCTGCCTTGAG	3800
GTTTTGGGTTGGGTCCATCAGCGATGATGCATCATGTGCTGCTCTTGGAG	3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCATG	3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCACCA	3950
TTTGGTATGTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA	4000
AAGGTAAACTGCTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT	4050
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC	4100
TATCAAGAGAAGCACAAAAGGCTCTAAGGAAATATACGTTGACAGGTACC	4150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATAGG	4200
TCTCTTAAACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG	4250
TCTCTTAAACATTGCTAATGTAGATGTTTATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTTCT	4300
ATGCTGCTGGAGTTAGAATATAATAGAATCTAACGGATACTTAAAGGT	4350
TCTTCTGAATCATGGAACTGTGCTGTTGATGGTCTAAATGAATG	4400
TCTTCTGAATCATGGAACTGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	

TGCCCTTAGTGCTCTTGGAGAGCTAATTAATCATCTGTCTAGGCTAAAGG	4450
TGTGTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4500
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4550
ATGGGGATATTTTTCCATACCAAGTTTACAGGGGTTGTCTCAGAATTGAT	4600
GGCCAGACGATGGTAAATCTTGAGATATTTAACAATAGCTGTGATGGTGG	4650
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA	4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
AGCCGATTTTTGTTTACCTATGTAGGGACCTTGTACAAATATCTTGATAA	4800
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC	. 4850
CACTCAAAGATGTAGAAAGCATCAATAAACGGCTTGATGTAGTTGAAGAA	4900
TTCACGGCAAACTCAGAAAGTATGCAAATCACTGGCCAGTATCTCCACAA	4950
ACTTCCAGACTTAGAAAGACTGCTCGGACGCATCAAGTCTAGCGTTCGAT	5000
CATCAGCCTCTGTGTGCCTGCTCTTCTGGGGAAAAAAGTGCTGAAACAA	50 50
CGAGTAAGTATCAATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT	5100
ATAGGACTAAAACATTACGGGTCTAGCTAAAGACTGTTCTCCTTCTTTTG	5150
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5200
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT	5250
GTTGGCTCTACAGAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT	5300
GTAAACTTCCTATATTAGTAGGAAAAAGCGGGCTAGAGTTATTTCTTTC	5350
CAATTCGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTGCCCAT	5400
CTATCTTTCATACTTTACAACAAAATGTCTGTCACTACTCAAAGCAATGC	5450
ATATGGCTTAGATCTCAACTCACACCCCGAGGATCCTAAAGGGATTTGCT	5500
TTTTATTCCTAATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT	5 550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAACGCTGAAACTC	5600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGAG	5650
GTCATTCACACCATAAGCTGCCTAGATGTCCTGAGATCTTTTGCAATCGC	5700
AGCAAGTCTCTCTGCTGGAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5750
CAGAAGCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC	5800
CAAGGACTATGGCATCCATTTGCAGTTGCAGCCGATGGTCAATTGCCTGT	5850
TCCGAATGATATACTCCTTGGCGAGGCTAGAAGAAGCAGTGGCAGCATTC	5900
ATCCTCGGTCATTGTTACTGACGGGACCAAACATGGGCGGAAAATCAACT	5950
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6000
CGTTAGATAATTACTCTATTCTTTGCAATCAGTTCTTCAACATGAATAAT	6050
AAATTCTGTTTTCTGTCTGCAGCTTGGCTGCTACGTGCCGTGTGAGTCTT	6100
GCGAAATCTCCCTCGTGGATACTATCTTCACAAGGCTTGGCGCATCTGAT	6150
AGAATCATGACAGGAGAGAGTAAGTTTTGTTCTCAAAATACCAATTCCTC	6200
GAACTATTTACTCAGATTTTGTCTGATTGGACAAGGTGGTTTTGCTTTTT	6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT	6350
AGTACTTTCGATGGATACGCCATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA	6450
ACTTATTGATTTATATCAGGTTTTTCGTCACCTGGTAGAGAAAGTTCAAT	6500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCACCCACGTGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG	6600

ATCTGATTATCAACCACGTGGTTGTGATCAAGACCTAGTGTTCTTGTACC	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	6700
ATGGCTGGAATACCAAACCAAGTGGTTGAAACAGCATCAGGTGCTGCTCA	6750
AGCCATGAAGAGATCAATTGGGGAAAACTTCAAGTCAAG	6800
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTCGCCCACAACAATGCCCCCATTGGCGAAGATGACTACGACAC	6900
TTTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	69 5 0
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAAATTGTTTCC	7050
TCGATTATAACAAGATTATATATGTATCTGTCGGTTTAGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAAATACTCACAAACAGTA	7150
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATAACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT	7250
TATAAGACTTAAGTTCTGAATTCCCTAAAACCAAAAAAAA	7300
TATTTTGTTGCATCTACAAACACACACACACACATAGTTTATAACTTAC	7350
TCATCACTGAGATTAACATCAGAATCATTCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTTCACGTAACC	7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCAACTCC	7500
AAATCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	,7600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG	7650
TACTCAACCTCATCCTCATCCTCCTCCTCTCAGAATCACCAGGACT	7700 ي
ATCCATCATCTCATCAAACCCATTAGACTTATCTAAATAAA	7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	-7800
GACTTGGGCGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA	_7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT	7900
ACTTCCTCTTCAAGTCATCATTCTCCTACATTGCGTCTCAGATTTC	7950
TCCATCCTTAGCTCCTCACTCACTTTCTCAGCTACTTCATTCCAATCCTC	8000
GTTCCTCAAACTCCTTCTACCCAATTGCAAAAACCTATCTCCCCAAACTT	8050
CAAGCAACACAA	8062
·	

Figure 11 (Continued)



Comments/References: AtMSH3 3' side antisense : AtMSH3 3' (13 = 2104bp) from pUC18/13 Sal1/Sst1/T4 into pCW164 BamH1/T4 in Agrobacterium LBA44O4

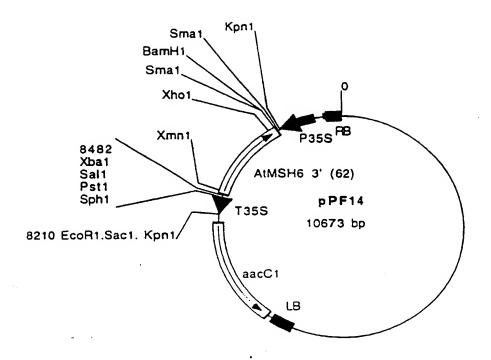


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sst1/T4 (1379bp) into pCW164 BamH1/T4

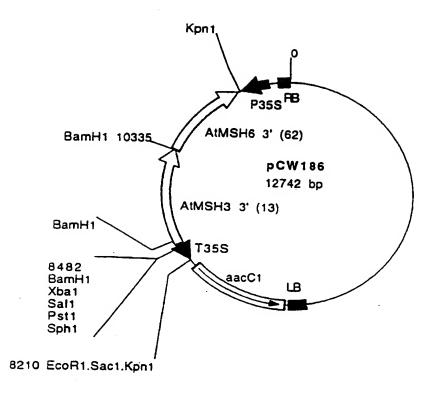
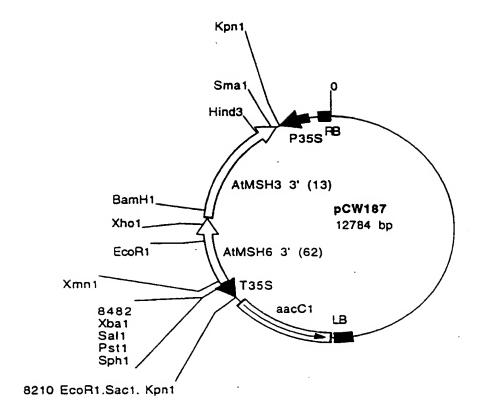


Figure 14

Comments/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (S8) 3' side (62=1379bp) Sal1/Sst1/T4 into pPF13 (pCW164 AtMSH3 (S5) 3' side (13=2104) antisens)/Sma1. in LBA4404



Comments/References: AtMSH3 3'/AtMSH6 3' antisens (D): AtMSH3 (S5) 3' side (13=2104bp) Sat1/Sst1/T4 into pPF14 (AtMSH6 (S8) 3'side (62=1379bp) antisense into pCW164)/Sma1. in LBA4404

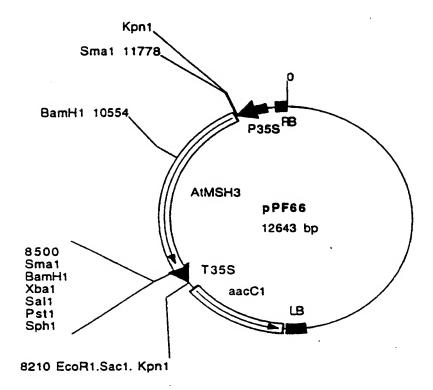
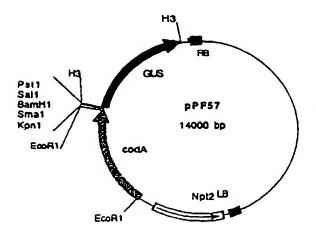


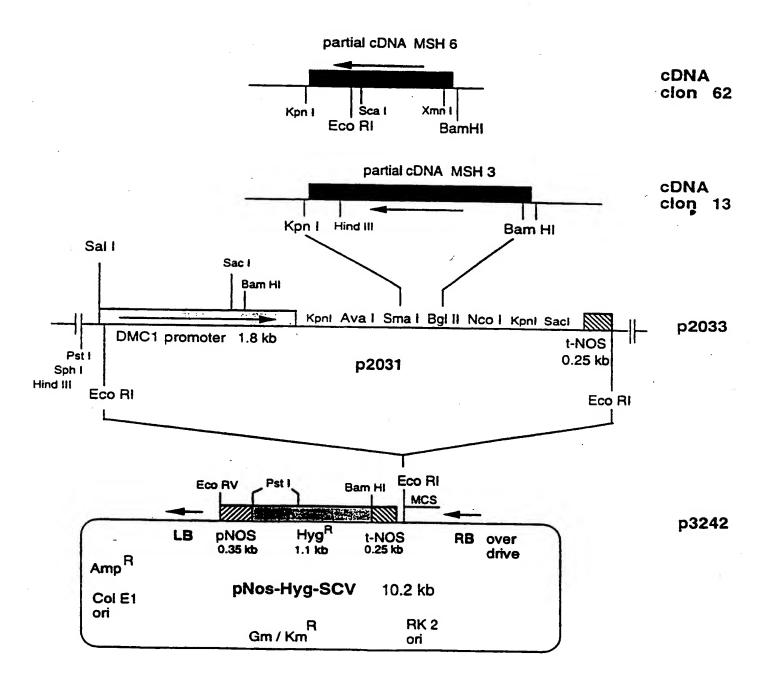
Figure 16

Comments/References: AtMSH3 (S8) complete, sense orientation : pPF26 (3342bp) Sma1 into pCW164 Sma1

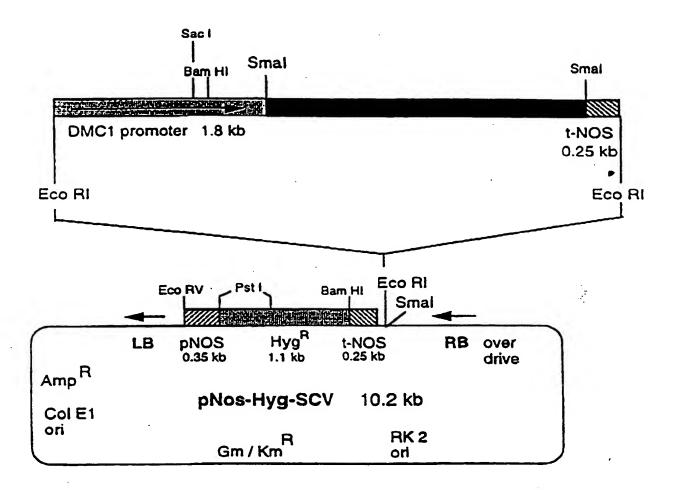


Comments/References: pPZP111 with codA EcoR1 cassette in EcoR1 site and HInd3 GUS cassette in Hind3 site. KanR. All genes under Promoter/terminator 35S

Figure 18



p3243



1

SEQUENCE LISTING

```
<110>
                Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
               Marie-Pascale; Freyssinet, Georges; Perez, Pascual.
               Methods for obtaining plant varieties
  <120>
  <130>
               395498C
  <150>
               PO9745
  <151>
               1997-10-10
  <160>
               98
  <210>
               1
  <211>
               23
  <212>
               DNA
 <213>
              Artificial sequence
 <220>
 <221>
              modified_base
 <222>
              11
 <223>
              I
 <220>
 <221>
              modified_base
 <222>
              14
 <223>
              I
 <220>
 <221>
              modified_base
 <222>
              17
 <223>
<220>
<223>
             Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
<300>
<301>
             Reenan and Kolodner
<302>
             Genetics
<303>
             132
<306>
             963-973
<307>
             1992
<400>
             1
ctggatccac nggnccnaay atg
                                                                        23
```

DOCID: <WO__9919492A2_I_>

<210>

<211>

<212>

2

23

DNA

مسرية فالفائد الداموه الدا

WO 99/19492 PCT/EP98/06977

2

```
Artificial sequence
 <213>
 <220>
              modified base
 <221>
 <222>
              15
 <223>
              Ι
 <220>
 <221>
              modified base
 <222>
              18
 <223>
              I
 <220>
              Degenerate oligonucleotides DOMU used to isolate AtMSH3 and
 <223>
              AtMSH6.
 <300>
 <301>
              Reenan and Kolodner
 <302>
              Genetics
 <303>
              132
<306>
              963-973
 <307>
              1992
 <400>
 ctggatccrt artgngtnrc raa
                                                                        23
 <210>
              3
 <211>
              24
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              MSH3 specific primer 636 for PCR using cDNA of Arabidopsis
              thaliana ecotype Columbia
 <400>
                                                                        24
 tgctagtgcc tcttgcaagc tcat
 <210>
 <211>
              27
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
 <223>
              Columbia containing adapter sequences ligated to both its
              ends
 <400>
```

3

ccatcctaat	acgactcact atagggc	
		27
<210>	5	
<211>	23	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Primer AP2 for PCR using cDNA of Arabidopsis thaliana e Columbia containing adapter sequences ligated to both i ends	ecotype its
<400>	5	
actcactata	gggctcgagc ggc	23
<210>	6	
<211>	30	
<212>	DNA	•
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer seas for popular	•
	MSH3 specific primer S525 for PCR using cDNA of Arabido thaliana ecotype Columbia	psis
<400>	6	
aggttctgat	tatgtgtgac gctttactta	30
<210>	7 .	
<211>	29	
<212>	DNA	
<213>	Artificial sequence	
		•
<220>		
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidop thaliana ecotype Columbia	sis
<400>	7	
ggatcgggta	tgggttttg agtgtgagg	29
<210>	8	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		•
<223>	MCV2 manifile anima and	
~463 7	MSH3 specific primer 635 for PCR using cDNA of Arabidops thaliana ecotype Columbia	sis

WO 99/19492 PCT/EP98/06977

< 400 >	8 .	
gcacgtgctt	gatggtgttt tcac	24
<210>	9	
<211>	28	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer S523 for PCR using cDNA of Arabidops thaliana ecotype Columbia	sis
<400>	9	
ccagacagca	tccagcatgg cagaagta	28
<210>	10	
<211>	33	
<212>	DNA	
<213>	Artificial sequence	
<220>	·	
<223>	MSH3 specific primer 1S5 for PCR using cDNA of Arabidops: thaliana ecotype Columbia	is
<400>	10	
atcccgggat	gggcaagcaa aagcagcaga cga	3 3
<210>	11	
<211>	27	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer S53 for PCR using cDNA of Arabidops thaliana ecotype Columbia	is
<400>	11	
gacaaagagc	gaaatgaggc cccttgg	27
<210>	12	-
<211>	1250	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<223>	Clone 52	

<400> 12

cccgggatgg	gcaagcaaaa	gcagcagacg	atttctcgtt	tcttcgctcc	caaacccaaa	60
tccccgactc	acgaaccgaa	tccggtagcc	gaatcatcaa	caccgccacc	gaagatatcc	120
gccactgtat	ccttctctcc	ttccaagcgt	aagcttctct	ccgaccacct	cgccgccgcg	180
tcacccaaaa	agcctaaact	ttctcctcac	actcaaaacc	cagtacccga	tcccaattta	240
caccaaagat	ttctccagag	atttctggaa	ccctcgccgg	aggaatatgt	tcccgaaacg	300
tcatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattcttcgg	agaagacgcg	420
gagatcgcag	cacgcgtgtt	gggtatttac	gctcatatgg	atcacaattt	catgacggcg	480
agtgtgccaa	catttcgatt	gaatttccat	gtgagaagac	tggtgaatgc	aggatacaag	540
attggtgtag	tgaagcagac	tgaaactgca	gccattaagt	cccatggtgc	aaaccggacc	6 00
ggcccttttt	tetggggaet	gccggcgctg	tataccaaag	ccacgcttga	agcggctgag	660
gatataagtg	gtggttgtgg	tggtgaagaa	ggttttggtt	cacagagtaa	tttcttggtt	720
tgtgttgtgg	atgagagagt	taagtcggag	acattaggct	gtggtattga	aatgagtttt	780
gatgttagag	tcggtgttgt	tggcgttgaa	atttcgacag	gtgaagttgt	ttatgaagag	840
ttcaatgata	atttcatgag	aagtggatta	gaggctgtga	ttttgagctt	gtcaccagct	900
gagctgttgc	ttggccagcc	tctttcacaa	caaactgaga	agtttttggt	ggcacatgct	960
ggacctacct	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	caatggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aaaatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	agctggaggc	tgctgaaaaa	ggaatgtctt	gcttgacagt	tcatacaatt	1140
atgaacatge	cacatctgac	tgttcaagcc	ctcgccctaa	cgttttgcca	tctcaaacag	1200
tttggatttg	aaaggatcct	ttaccaaggg	gcctcatttc	gctctttgtc		1,250

<210> 13
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> MSH3 specific primer 2S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<4005	13	
accegggte	aaaatgaaca agttggtttt agtc	34
<210>	14	
<211>	27	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer S52 for PCR using cDNA of Arabic	dopsis
	thaliana ecotype Columbia	
<400>	14	
gccacatctg	actgttcaag ccctcgc	27
<210>	15	
<211>	2110	•
<212>	DNA	•
<213>	Arabidopsis thaliana ecotype Columbia	
<223>	Clone 13	· fr
<400>	15	· 4
gccacatctg	actgttcaag ccctcgccct aacgttttgc catctcaaac agtttggatt	r. 60
tgaaaggatc	ctttaccaag gggcctcatt tcgctctttg tcaagtaaca cagagatgac	120
teteteagee	aatactctgc aacagttgga ggttgtgaaa aataattcag atggatcgga	180
atctggctcc	ttattccata atatgaatca cacacttaca gtatatggtt ccaggcttct	240
tagacactgg	gtgactcatc ctctatgcga tagaaatttg atatctgctc ggcttgatgc	300
tgtttctgag	atttctgctt gcatgggatc tcatagttct tcccagctca gcagtgagtt	360
ggttgaagaa	ggttctgaga gagcaattgt atcacctgag ttttatctcg tgctctcctc	
agtettgaca	gctatgtcta gatcatctga tattcaacgt ggaataacaa gaatctttca	
teggaetget	aaagccacag agttcattgc agttatggaa gctattttac ttgcggggaa	•
gcaaattcag	cggcttggca taaagcaaga ctctgaaatg aggagtatgc aatctgcaac	600
tgtgcgatct	actettttga gaaaattgat ttetgttatt teateceetg ttgtggttga	
caatgccgga	aaacttctct ctgccctaaa taaggaagcg gctgttcgag gtgacttgct	
cgacatacta	atcacttcca gcgaccaatt tcctgagctt gctgaagctc gccaagcagt	780
tttagtcatc	agggaaaagc tggattcctc gatagcttca tttcgcaaga agctcgctat	840

togaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
ttccaaggtc	cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	ttcgatatca	960
tcccccagaa	atagtagctg	gcttggatga	gctagctcta	gcaactgaac	atcttgccat	1020
tgtgaaccga	gcttcgtggg	atagtttcct	caagagtttc	agtagatact	acacagattt	1080
taaggctgcc	gttcaagctc	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag	aactatgtcc	gtcccgagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtct	ggtcgtcatc	ctgtactgga	gactatatta	caagataact	tcgtcccaaa	1260
tgacacaatt	ttgcatgcag	aaggggaata	ttgccaaatt	atcaccggac	ctaacatggg	1320
aggaaagagc	tgctatatcc	gtcaagttgc	tttaatttcc	ataatggctc	aggttggttc	1380
ctttgtacca	gcgtcattcg	ccaagctgca	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgcttcagac	agtatccagc	atggcagaag	tacctttcta	gaagaattaa	gtgaagcgtc	1500
acacataatc	agaacctgtt	cttctcgttc	gcttgttata	ttagatgagc	ttggaagagg	1560
cactagcaca	cacgacggtg	tagccattgc	ctatgcaaca	ttacagcatc	tcctagcaga	1620
aaagagatgt	ttggttcttt	ttgtcacgca	ttaccctgaa	atagctgaga	tcagtaacgg	1680
attcccaggt	tctgttggga	cataccatgt	ctcgtatctg	acattgcaga	aggataaagg	1740
cagttatgat	catgatgatg	tgacctacct	atataagctt	gtgcgtggtc	tttgcagcag	1800
gagctttggt	tttaaggttg	ctcagcttgc	ccagatacct	ccatcatgta	tacgtcgagc	1860
catttcaatg	gctgcaaaat	tggaagctga	ggtacgtgca	agagagaa	atacacgcat	1920
gggagaacca	gaaggacatg	aagaaccgag	aggcgcagaa	gaatctattt	cggctctagg	1980
tgacttgttt	gcagacctga	aatttgctct	ctctgaagag	gacccttgga	aagcattcga	2040
gtttttaaag	catgcttgga	agattgctgg	caaaatcaga	ctaaaaccaa	cttgttcatt	2100
ttgacccggg						2110

<210>	16
<211>	29
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	16

ggat	cggg	rta c	tggg	ttt	g ag	tgtg	agg									2.9
<210<211<212<213	> >		17 30 DNA		ial	cagu	anca							٠		
			AIC	. 1 1 1 C	ıaı	sequ	ence	•								
<220 <223					ecif a ec					or P	CR u	sing	CDN	A of	Arab	oidopsis
< 400	>		17													
aggt	tctg	at t	atgt	gtga	ic go	ttta	ctta	ı								3 (
<210 <211			18 352	2												
<212	>		DNA													
<213	>		Ara	bido	psis	: tha	iia.	a ec	otyp	e Co	lumi:	oia				•
<220	>															*···
<221	. >		CDS	;												~
<222	>		(10	0)	(3	342)										**
<223	>	•			full tide		gth	CDNA	and	l ded	luced	sec	luenc	e of	the	encoded
< 400	>		18						•							·*
ccta	agaa	ag d	gcgo	gaaa	a tt	ggca	acco	aag	jttc:	gcca	tago	cac	gac o	acga	acctt	60
catt	tcto	ett a	aacç	gagg	ga ga	ittac	gaat	aaa	agcaa	ict						99
														ccc Pro 15		147
			-				_		_	_	-	_		tca Ser		195
_		_	_			_		_						aag Lys	_	243
_				-			_	_					_	cct Pro		291
								_		_				cac His		339

aga Arg	t t t Phe	ctc Leu	cag Gln	aga Arg 85	Pue	ctg Leu	gaa Glu	CCC Pro	tcg Ser	Pro	gag Glu	gaa Glu	tat Tyr	gtt Val	ccc	387
gaa Glu	acg Thr	cca Ser	cca Ser 100	ser	agg Arg	aaa Lys	cac Tyr	aca Thr	Pro	ttg Leu	gaa Glu	cag Gln	caa Gln 110	gtg Val	gtg Val	435
gag Glu	cta Leu	aag Lys 115	agc Ser	aag Lys	tac Tyr	cca Pro	gat Asp 120	gtg Val	gtt Val	ttg Leu	atg Met	gtg Val 125	gaa Glu	gtt Val	ggt Gly	483
tac Tyr	agg Arg 130	tac Tyr	aga Arg	ttc Phe	ttc Phe	gga Gly 135	gaa Glu	gac Asp	gcg Ala	gag Glu	atc Ile 140	gca Ala	gca Ala	cgc	gtg Val	531
ttg Leu 145	ggt Gly	att Ile	tac Tyr	gct Ala	cat His 150	atg Met	gat Asp	cac His	aat Asn	ttc Phe 155	atg Met	acg Thr	gcg Ala	agt Ser	gtg Val 160	579
cca Pro	aca Thr	ttt Phe	cga Arg	ttg Leu 165	aat Asn	ttc Phe	cat His	gtg Val	aga Arg 170	aga Arg	ctg Leu	gtg Val	aat Asn	gca Ala 175	gga Gly	627
tac Tyr	aag Lys	att Ile	ggt Gly 180	gta Val	gcg Val	aag Lys	cag Gln	act Thr 185	gaa Glu	act Thr	gca Ala	gcc Ala	att Ile 190	aag Lys	tcc Ser	675
cat His	ggt Gly	gca Ala 195	aac Asn	cgg Arg	acc Thr	ggc	cct Pro 200	ttt Phe	ttc Phe	cgg Arg	gga Gly	ctg Leu 205	tcg Ser	gcg Ala	ttg Leu	723
tat Tyr	acc Thr 210	aaa Lys	gcc Ala	acg Thr	ctt Leu	gaa Glu 215	gcg Ala	gct Ala	gag Glu	gat Asp	ata Ile 220	agt Ser	ggt Gly	ggt Gly	tgt Cys	771
ggt Gly 225	ggt Gly	gaa Glu	gaa Glu	ggt Gly	ttt Phe 230	ggt Gly	tca Ser	cag Gln	agt Ser	aat Asn 235	ttc Phe	ttg Leu	gtt Val	tgt Cys	gtt Val 240	819
gtg Val	gat Asp	gag Glu	aga Arg	gtt Val 245	aag Lys	tcg Ser	gag Glu	aca Thr	tta Leu 250	ggc Gly	tgt Cys	ggt Gly	att Ile	gaa Glu 255	atg Met	867
agt Ser	ttt Phe	gat Asp	gtt Val 260	aga Arg	gtc Val	ggt Gly	gtt Val	gtt Val 265	ggc Gly	gtt Val	gaa Glu	att Ile	tcg Ser 270	aca Thr	ggt Gly	915
gaa Slu	vaı	gtt Val 275	tat Tyr	gaa Glu	gag Glu	Phe	aat Asn 280	gat Asp	aat Asn	ttc Phe	atg Met	aga Arg 285	agt Ser	gga Gly	tta Leu	963

													ggc Gly		1011
													gga Gly		1059
			-		-			_	_	_	-		agc Ser 335		1107
		-	_	_		_				_			agc Ser	-	1155
			_	_	_				-		_	-	gaa Glu		1203
													cat His		125 🖁
													ttt Phe		1299
	_							_		_		_	tca Ser 415	-	1347
													gag Glu		1395
					_		_	-					cat His		1443
_						-		_		•		_	cac His	tgg Trp	. 1491
										Ser			ctt Leu		1539
_	_					- ,	_	_	 Ser		_		tcc Ser 495	Gln	1587

										11							
				5,00					505	061	GIU	Arg	gca Ala	Ile 510	Val.	Ser	1635
			515	-			9	520		vaı	red	inr	gct Ala 525	Met	Ser	Arg	1683
	5	30				9	535			ALG	116	540	cat His	Arg	Thr	Ala	1731
54	5				!	550				GIU	555	iie	tta (Leu .	Ala	Gly 560	1779
					565		- 7 -		Jys	570	Asp	ser	gaa a Glu N	let 1	Arg	Ser	1827
			5	80			-9 J	5	85	Leu	Leu /	Arg :	_	eu I 90	le s	Ser	1875
		5	95				60	0	.sp .	nsii i	ara C	- 3∓À-1	aaa c Lys L 605	eu L	eu s	Ser	1923
	61	0				61	.5	ı A	rg (ala t	ksp L 6	leu I 20	etc g.	sp I	le L	eu	1971
625					63	0		.	.u.r	6 eu	35	Iu A	rct co	rg G	ln A 6	la 40	2019
				64	45	- - -	ם בפי	a As	6 6	er s 50	er I	le A	ct to la Se	r Ph	ie Ai	rg	2067
			66	0		J		66	5	ne r	eu G	in Va	tg tc al Se 67	r Gl O	y II	Le	2115
		67	5				680	A.S	Ď 26	ar r)	⁄s va	11 Pr 68		s As:	n Tr	P	2163
grg Val	aaa Lys 690	gta Val	a aat L Asi	ag 1 Se	c acc	aag Lys 695	wy S	ac Th	t at	c cg e Ar	a ta g Ty 70	r Hi	t cc	e cca	a ga o Gl	a u	2211

ata Ile 705	gta Val	gct Ala	ggc Gly	ttg Leu	gat Asp 710	gag Glu	cta Leu	gct Ala	cta Leu	gca Ala 715	act Thr	gaa Glu	cat His	ctt Leu	gcc Ala 720	2259
att Ile	gtg Val	aac Asn	cga Arg	gct Ala 725	tcg Ser	tgg Trp	gat Asp	agt Ser	ttc Phe 730	ctc Leu	aag Lys	agt Ser	ttc Phe	agt Ser 735	aga Arg	2307
tac Tyr	tac Tyr	aca Thr	gat Asp 740	ttt Phe	aag Lys	gct Ala	gcc Ala	gtt Val 745	caa Gln	gct Ala	ctt Leu	gct Ala	gca Ala 750	ctg Leu	gac Asp	2355
														gtc Val		2403
														cag Gln		2451
Gly 785	Arg	His	Pro	Val	Leu 790	Glu	Thr	Ile	Leu	Gln 795	ąsk	Asn	Phe	gtc Val	Pro 800	249\$
Asn	Asp	Thr	Iļe	Leu 805	His	Ala	Glu	Gly	Glu 810	Tyr	Cys	Gln	Ile	atc Ile 815	Thr	2547
Gly	Pro	Asn	Met 820	Gly	Gly	Lys	Ser	Cys 825	Tyr	Ile	Arg	Gln	Val 830	gct Ala	Leu	2595
Ile	Ser	Ile 835	Met	Ala	Gln	Val	Gly 840	Ser	Phe	Val	Pro	Ala 845	Ser	ttc Phe	Ala	2643
Lys	Leu 850	His	Val	Leu	Asp	Gly 855	Val	Phe	Thr	Arg	Met 860	Gly	Ala	tca Ser	Asp	2691
														gaa Glu		2739
														tta Leu 895		2787
														gcc Ala		2835

gca Ala	aca Thr	tta Leu 915	cag Gln	cat His	ctc Leu	cta Leu	gca Ala 920	gaa Glu	aag Lys	aga Arg	tgt Cys	ttg Leu 925	gtt Val	ctt Leu	ttt Phe	2883
gtc Val	acg Thr 930	cat His	tac Tyr	cct Pro	gaa Glu	ata Ile 935	gct Ala	gag Glu	atc Ile	agt Ser	aac Asn 940	gga Gly	ttc Phe	cca Pro	ggt Gly	2931
tct Ser 945	gtt Val	ggg	aca Thr	tac Tyr	cat His 950	gtc Val	tcg Ser	tat Tyr	ctg Leu	aca Thr 955	ttg Leu	cag Gln	aag Lys	gat Asp	aaa Lys 960	2979
Gly	agt Ser	tat Tyr	gat Asp	cat His 965	gat Asp	gat Asp	gtg Val	acc Thr	tac Tyr 970	cta Leu	tat Tyr	aag Lys	ctt Leu	gtg Val 975	cgt Arg	3027
ggt Gly	ctt Leu	tgc Cys	agc Ser 980	agg Arg	agc Ser	ttt Phe	ggt Gly	ttt Phe 985	aag Lys	gtt Val	gct Ala	cag Gln	ctt Leu 990	gcc Ala	cag Gln	3075
ata Ile	cct Pro	cca Pro 995	tca Ser	tgt Cys	ata Ile	Arg	cga Arg	gcc Ala	att Ile	tca Ser	Met	gct Ala .005	gca Ala	aaa Lys	ttg Leu	3123
GIU	gct Ala 010	gag Glu	gta Val	cgt Arg	Ala	aga Arg .015	gag Glu	aga Arg	aat Asn	Thr	cgc Arg 1020	atg Met	gga Gly	gaa Glu	cca Pro	3171
gaa Glu 1025	GIY	cat His	gaa Glu	Glu	ccg Pro 1030	aga Arg	ggc Gly	gca Ala	Glu	gaa Glu 035	tct Ser	att Ile	tcg Ser	gct Ala 1	cta Leu .040	3219
ggt Gly	gac Asp	ttg Leu	Phe	gca Ala 045	gac Asp	ctg Leu	aaa Lys	Phe	gct Ala 050	ctc Leu	tct Ser	gaa Glu	Glu	gac Asp .055	cct Pro	3267
tgg Trp	aaa Lys	Ala	ttc Phe 060	gag Glu	ttt Phe	tta Leu	Lys	cat His 065	gct Ala	tgg Trp	aag Lys	Ile	gct Ala 070	Gly	a aa Lys	3315
atc Ile	Arg	cta Leu 075	aaa Lys	cca Pro	act Thr	Cys	tca Ser 080	ttt Phe	tgat	ttaa	țc t	taac	atta	t		3362
agca	actg	ca a	ggtc	ttga	t ca	tctg	ttag	ttg	cgta	cta	actt	atgt	gt a	ttag	tataa	3422
caag	aaaa	ga g	aatt	agag	a ga	tgga	ttct	aat	ccgg	tgt	tgca	gtac	at c	ttt	ctcca	3482
cccg	cata	aa a	aaaa	aaaa	a aa	aaaa	aaaa	aaa	aaaa	aaa						3522
<210	>		19				•									
<211			108	1												

ISDOCID: <WO__9919492A2_I_>

<212>

PRT

and the state of t

WO 99/19492 PCT/EP98/06977

<213 <223						tha MSH		a ec	otyp	e Co	lumb	ia			
< 400) >		19												
Met 1	Gly	Lys	Gln	Lys 5	Gln	Gln	Thr	Ile	Ser 10	Arg	Phe	Phe	Ala	Pro 15	Lys
Pro	Lys	Ser	Pro 20	Thr	His	Glu	Pro	Asn 25	Pro	Val	Ala	Glu	Ser 30	Ser	Thr
Pro		Pro 35	Lys	Ile	Ser	Ala	Thr 40	Val	Ser	Phe	Ser	Pro 45	Ser	Lys	Arg
Lys	Leu 50	Leu	Ser	Asp	His	Leu 55	Ala	Ala	Ala	Ser	Pro 60	Lys	Lys	Pro	Lys
Leu 65	Ser	Pro	His	Thr	Gln 70	Asn	Pro	Val	Pro	Asp 75	Pro	Asn	Leu	His	Gln 80
Arg	Phe	Leu	Gln	Arg 85	Phe	Leu	Glu	Pro	Ser 90	Pro	Glu	Glu	Tyr	Val 95	Pro
Glu	Thr	Ser	Ser 100	Ser	Arg	Lys	Tyr	Thr 105	Pro	Leu	Glu	Gln	Gln 110	Val	Val
Glu	Leu	Lys 115	Ser	Lys	Tyr	Pro	Asp 120	Val	V al	Leu	Met	Val 125	Glu	Val	Gly
Tyr	Arg 130	Tyr	Arg	Phe	Phe	Gly 135	Glu	Asp	Ala	Glu	Ile 140	Ala	Ala	Arg	Val
Leu 145	Gly	Ile	Tyr	Ala	His 150	Met	Asp	His	Asn	Phe 155	Met	Thr	Ala	Ser	Val 160
Pro	Thr	Phe	Arg	Leu 165		Phe	His	Val	Arg 170		Leu	Val	Asn	Ala 175	Gly
Tyr	Lys	Ile	Gly 180	Val	Val	Lys	Gln	Thr 185	Glu	Thr	Ala	Ala	Ile 190	Lys	Ser
His	Gly	Ala 195	Asn	Arg	Thr	Gly	Pro 200	Phe	Phe	Arg	Gly	Leu 205	Ser	Ala	Leu
Tyr	Thr 210	Lys	Ala	Thr	Leu	Glu 215		Ala	Glu	Asp	Ile 220	Ser	Gly	Gly	Суѕ
Gly 225	Gly	Glu	Glu	Gly	Phe 230	-	Ser	Gln	Ser	Asn 235		Leu	Val	Cys	Val 240
Val	Asp	Glu	Arg	Val 245		Ser	Glu	Thr	Leu 250		Cys	Gly	Ile	Glu 255	Met

Ser	Phe	Asp	Val 260	Arg	Val	Gly	Val	Val 265	Gly	Val	Glu	Ile	Ser 270	Thr	Gly
Glu	Val	Val 275	Tyr	Glu	Glu	Phe	Asn 280	Asp	Asn	Phe	Met	Arg 285	Ser	Gly	Leu
Glu _,	Ala 290	Val	Ile	Leu	Ser	Leu 295	Ser	Pro	Ala	Glu	Leu 300	Leu	Leu	Gly	Gln
Pro 305	Leu	Ser	Gln	Gln	Thr 310	Glu	Lys	Phe	Leu	Val 315	Ala	Met	Ala	Gly	Pro 320
Thr	Ser	Asn	Val	Arg 325	Val	Glu	Arg	Ala	Ser 330	Leu	Asp	Cys	Phe	Ser 335	Asn
Gly	Asn	Ala	Val 340	Asp	Glu	Val	Ile	Ser 345	Leu	Cys	Glu	Lys	Ile 350	Ser	Ala
Gly	Asn	Leu 355	Glu	Asp	Asp	Lys	Glu 360	Met	Lys	Leu	Glu	Ala 365	Ala	Glu	Lys
Gly	Met 370	Ser	CÀR	Leu	Thr	Val 375	His	Thr	Ile	Met	Asn 380	Met	Pro	His	Leu
Thr 385	Val	Gln	Ala	Leu	Ala 390	Leu	Thr	Phe	Cys	His 395	Leu	Lys	Gln	Phe	Gly 400
Phe	Glu	Arg	Ile	Leu 405	Tyr	Gln	Gly	Ala	Ser 410	Phe	Arg	Ser	Leu	Ser 415	Ser
Asn	Thr	Glu	Met 420	Thr	Leu	Ser	Ala	Asn 425	Thr	Leu	Gln	Gln	Leu 430	Glu	Val
Val	Lys	Asn 435	Asn	Ser	Asp	Gly	Ser 440	Glu	Ser	Gly	Ser	Leu 445	Phe	His	Asn
Met	Asn 450	His	Thr	Leu	Thr	Val 455	Tyr	Gly	Ser	Arg	Leu 460	Leu	Arg	His	Trp
Val 465	Thr	His	Pro	Leu	Cys 470	Asp	Arg	Asn	Leu	Ile 475	Ser	Ala	Arg	Leu	Asp 480
Ala	Val	Ser	Glu	Ile 485	Ser	Ala	Cys	Met	Gly 490	Ser	His	Ser	Ser	Ser 495	Gln
Leu	Ser	Ser	Glu 500	Leu	Val	Glu	Glu	Gly 505	Ser	Glu	Arg	Ala	Ile 510	Val	Ser
Pro	Glu	Phe 515	туr	Leu	Val	Leu	Ser 520	Ser	Val	Leu	Thr	Ala 525	Met	Ser	Arg
Ser	Ser 530	Asp	Ile	Gln	Arg	Gly 535	Ile	Thr	Arg	Ile	Phe 540	His	Arg	Thr	Ala

Lys 545		Thr	Glu	Phe	Ile 550	Ala	Val	Met	Glu	Ala 555	Ile	Leu	Leu	Ala	Gly 560
Lys	Gln	Ile	Gln	Arg 565	Leu	Gly	Ile	Lys	Gln 570	Asp	Ser	Glu	Met	Arg 575	Ser
Met	Gln	Ser	Ala 580	Thr	Val	Arg	Ser	Thr 585	Leu	Leu	Arg	Lys	Leu 590	Ile	Ser
Val	Ile	Ser 595	Ser	Pro	Val	Val	Val 600	Asp	Asn	Ala	Gly	Lys 605	Ļeu	Leu	Ser
Ala	Leu 610	Asn	Lys	Glu	Ala	Ala 615	Val	Arg	Gly	Asp	Leu 620	Leu	Asp	Ile	Leu
Ile 625	Thr	Ser	Ser	Asp	Gln 630	Phe	Pro	Glu	Leu	Ala 635	Glu	Ala	Arg	Gln	Ala 640
Val	Leu	Val	Ile	Arg 645	Glu	Lys	Leu	Asp	Ser 650	Ser	Ile	Ala	Ser	Phe 655	Arg
Lys	Lys	Leu	Ala 660	Ile	Arg	Asn	Leu	Glu 665	Phe	Leu	Gln	Val	Ser 670	Gly	Ile
Thr	His	Leu 675	Ile	Glu	Leu	Pro	Val 680	Asp	Ser	Lys	Val	Pro 685	His	Asn	Trp
Val	Lys 690	Val	Asn	Ser	Thr	Lys 695	Lys	Thr	Ile	Arg	Tyr 700	His	Pro	Pro	Glu
Ile 705	Val	Ala	Gly	Leu	Asp 710	Glu	Leu	Ala	Leu	Ala 715	Thr	Glu	His	Leu	Ala 720
Ile	Val	Asn	Arg	Ala 725	Ser	Trp	Asp	Ser	Phe 730	Leu	Lys	Ser	Phe	Ser 735	Arg
Tyr	Tyr	Thr	Asp 740	Phe ⁻	Lys	Ala	Ala	Val 745	Gln	Äla	Leu	Ala	Ala 750	Leu	Asp
Cys	Leu	His 755	Ser	Leu	Ser	Thr	Leu 760	Ser	Arg	Asn	Lys	Asn 765	Tyr	Val	Arg
Pro	Glu 770	Phe	Val	Asp	Asp	Cys 775	Glu	Pro	Val	Glu	Ile 780	Asn	Ile	Gln	Ser
Gly 785	Arg	His	Pro	Val	Leu 790	Glu	Thr	Ile	Leu	Gln 795	Asp	Asn	Phe	Val	Pro 800
Asn	Asp	Thr	Ile	Leu 805	His	Ala	Glu	Gly	Glu 810	Tyr	Cys	Gln	Ile	Ile 815	Thr
Gly	Pro	Asn	Met 820	Gly	Gly	Lys	Sar	Cys 825	Tyr	Ile	Arg	Gln	Val 830	Ala	Leu

Ile Ser Ile Met Ala Gln Val Gly Ser Phe Val Pro Ala Ser Phe Ala 835 840 845

Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp

50 855 860

Ser Ile Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala 865 870 875 880

Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp 885 890 895

Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr 900 905 910

Ala Thr Leu Gln His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe 915 920 925

Val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Asn Gly Phe Pro Gly 930 935 940

Ser Val Gly Thr Tyr His Val Ser Tyr Leu Thr Leu Gln Lys Asp Lys 945 950 950 955

Gly Ser Tyr Asp His Asp Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg 965 970 975

Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln 980 985 990

Ile Pro Pro Ser Cys Ile Arg Arg Ala Ile Ser Met Ala Ala Lys Leu 995 1000 1005

Glu Ala Glu Val Arg Ala Arg Glu Arg Asn Thr Arg Met Gly Glu Pro 1010 1015 1020

Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser Ile Ser Ala Leu 1025 1030 1035 1040

Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Leu Ser Glu Glu Asp Pro 1045 1050 1055

Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys
1060 1065 1070

Ile Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080

<210> 20 <211> 24 <212> DNA

<213> Artificial sequence

<220>		
<223>	MSH6 specific primer 638 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia	5
<400>	20	
tctctaccag	gggacgaaaa accg	24
<210>	21	
<211>	28	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Primer S81 for PCR using cDNA of Arabidopsis thaliana ecot Columbia	- Ab
<400>	21	
cgtcgccttt	agcatecet teetteae	28
<210>	···	
<210>	22 30	
<212>	DNA	
<213>	Artificial sequence	
<220>	a a	
<223>	MSH6 specific primer S823 for PCR using cDNA of Arabidops: thaliana ecotype Columbia	is
<400>	22	
gcttggcgca	a tctaatagaa tcatgacagg	30
<210>	23	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer 637 for PCR using cDNA of Arabidopsisthaliana ecotype Columbia	s
<400>	23	
gacagcgtca	a gttcttcaga atgc	24
<210>	24	
<211>	33	
<212>	DNA	

<213>	Artificial sequence	
<220 > <223 >	MSH6 specific primer 1S8 for PCR using cDNA of Arabidops: thaliana ecotype Columbia	is
<400>	24	
atcccgggat	gcagegeeag agategattt tgt	3:
<210>		
<211>	27	
<212>	DNA	
<213>		
(213)	Artificial sequence	
<220>		
<223>	MSH6 specific primer S83 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	s
<400>	25	•
cgctatctat	ggctgcttcg aattgag	27
<210>	26	
<211>	1385	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<223>	Clone 43	
<400>	26	
cccgggatgc	agcgccagag atcgattttg tetttettee aaaaacccae ggcggcgact	60
acgaagggtt	tggtttccgg cgatgctgct agcggcgggg gcggcagcgg aggaccacga	120
tttaatgtga	aggaagggga tgctaaaggc gacgcttctg tacgttttgc tgtttcgaaa	180
tctgtcgatg	aggttagagg aacggatact ccaccggaga aggttccgcg tcgtgtcctg	240
ccgtctggat	ttaagccggc tgaatccgcc ggtgatgctt cgtccctgtt ctccaatatt	300
atgcataagt	ttgtaaaagt cgatgatcga gattgttctg gagagaggag ccgagaagat	360
gttgttccgc	tgaatgattc atctctatgt atgaaggcta atgatgttat tcctcaattt	420
cgttccaata	atggtaaaac tcaagaaaga aaccatgctt ttagtttcag tgggagagct	480
gaacttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
gggatgcgtc	cacgtgcttc tcgcttgaag cgagttctgg aggatgaaat gacttttaag	600
gaggataagg	ttcctgtatt ggactctaac aaaaggctga aaatgctcca ggatccggtt	660

WO 99/19492

cgcggagaga	agaaagaagt	aaacgaagga	accaaatttg	aatggcttga	gtcttctcga	720
atcagggatg	ccaatagaag	acgtcctgat	gatccccttt	acgatagaaa	gaccttacac	780
ataccacctg	atgttttcaa	gaaaatgtct	gcatcacaaa	agcaatattg	gagtgttaag	840
agtgaatata	tggacattgt	gcttttcttt	aaagtgggga	aattttatga	gctgtatgag	900
ctagatgcgg	aattaggtca	caaggagctt	gactggaaga	tgaccatgag	tggtgtggga	950
aaatgcagac	aggttggtat	ctctgaaagt	gggatagatg	aggcagtgca	aaagctatta	1020
gctcgtggat	ataaagttgg	acgaatcgag	cagctagaaa	catctgacca	agcaaaagcc	1080
agaggtgcta	atactataat	tccaaggaag	ctagttcagg	tattaactcc	atcaacagca	1140
agcgagggaa	acatcgggcc	tgatgccgtc	catcttcttg	ctataaaaga	gatcaaaatg	1200
gagctacaaa	agtgttcaac	tgtgtatgga	tttgcttttg	ttgactgtgc	tgccttgagg	1250
ttttgggttg	ggtccatcag	cgatgatgca	tcatgtgctg	ctcttggagc	gttattgatg	1320
caggilicic	caaaggaagt	gttatatgac	agtaaagggc	tatcaagaga	agcacaaaag	1330
gctctaagga	aatatacgtt	gacagggtct	acggcggtac	agttggctcc	agtaccacaa	1440
gtaatggggg	atacagatgc	tgctggagtt	agaaatataa	tagaatctaa	cggatacttt	1500
aaaggttctt	ctgaatcatg	gaactgtgct	gttgatggtc	taaatgaatg	tgatgttgcc	1560
cttagtgctc	ttggagagct	aattaatcat	ctgtctaggc	taaagctaga	agatgtactt	1620
aagcatgggg	atattttcc	ataccaagtt	tacaggggtt	gtctcagaat	tgatggccag	1680
acgatggtaa	atcttgagat	atttaacaat	agctgtgatg	gtggtccttc	agggaccttg	1740
tacaaatatc	ttgataactg	tgttagtcca	actggtaagc	gactcttaag	gaattggatc	1800
tgccatccac	tcaaagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaattc	1860
acggcaaact	cagaaagtat	gcaaatcact	ggccagtatc	tccacaaact	tccagactta	1920
gaaagactgc	tcggacgcat	caagtctagc	gttcgatcat	cagcctctgt	gttgcctgct	1980
cttctgggga	aaaaagtgct	gaaacaacga	gttaaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatct	gttgttggct	ctacagaagg	aatcaaatat	gatgagtttg	2100
ctttataaac	tctgtaaact	tcctatatta	gtaggaaaaa	gcgggctaga	gttatttctt	2160
tctcaattcg	aagcagccat	agatagcg				2188

<210> <211> <212> <213> <223> <400>	27 1385 DNA Arabidop Clone 62	osis chalian	aa ecotype (Columbia		
catcagecto	tgtgttgcct	gctcttctgg	ggaaaaaagt	: gctgaaacaa	cgagttaaag	60
catttgggca	aattgtgaaa	gggttcagaa	gtggaattga	tctgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	aactctgtaa	acttcctata	ttagtaggaa	180
aaagcgggct	agagttattt	ctttctcaat	tcgaagcagc	catagatago	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgctgaaac	tctcacaata	cttatcgaac	300
tttttatcga	aagagcaact	caatggtctg	aggtcattca	caccataago	tgcctagatg	360
tcctgagatc	ttttgcaatc	gcagcaagtc	tctctgctgg	aagcatggco	aggcctgtta	420
ttttcccga	atcagaagct	acagatcaga	accagaaaac	aaaagggcca	atacttaaaa	480
tccaaggact	atggcatcca	tttgcagttg	cagccgatgg	tcaattgcct	gttccgaatg	540
atatactcct	tggcgaggct	agaagaagca	gtggcagcat	tcatcctcgg	tcattgttac	600
gacgggacc	aaacatgggc	ggaaaatcaa	ctcttcttcg	tgcaacatgt	ctggccgtta	660
ctttgccca	acttggctgc	tacgtgccgt	gtgagtcttg	cgaaatctcc	ctcgtggata	720
tatcttcac	aaggcttggc	gcatctgata	gaatcatgac	aggagagagt	acctttttgg	780
agaatgcac	tgagacagcg	tcagttcttc	agaatgcaac	tcaggattca	ctagtaatcc	840
tgacgaact	gggcagagga	actagtactt	tcgatggata	cgccattgca	tactcggttt	900
tegteacet	ggtagagaaa	gttcaatgtc	ggatgctctt	tgcaacacat	taccaccctc	960
caccaagga	attegegtet	cacccacgtg	tcacctcgaa	acacatggct	tgcgcattca	1020
atcaagatc	tgattatcaa	ccacgtggtt	gtgatcaaga	cctagtgttc	ttgtaccgtt	1080
aaccgaggg	agcttgtcct	gagagctacg	gacttcaagt	ggcactcatg	gctggaatac	1140
aaaccaagt	ggttgaaaca	gcatcaggtg	ctgctcaagc	catgaagaga	tcaattgggg	1200
aaacttcaa	gtcaagtgag	ctaagatctg	agttctcaag	tctgcatgaa	gactggctca	1260
gtcattggt	gggtatttct	cgagtcgccc	acaacaatgc	cccattggc	gaagatgact	1320
cgacacttt	gttttgctt.i	tggcatgaga	tcaaatcctc	ttactgtgtt	cccaaataac	1380

ccggg	1385
<210>	28
<211>	34
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH6 specific primer 2S8 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	26
atcccgggtt	atttgggaac acagtaagag gatt 34
<210>	29
<211>	27
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH6 specific primer S82 for PCR using cDNA of Arabidopsis
(223)	thaliana ecotype Columbia
	· · · · · · · · · · · · · · · · · · ·
<400>	29
gcgttcgatc	atcagcctct gtgttgc 27
<210>	30
<211>	3606
<212>	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<220>	
<221>	CDS
<222>	(142)(3468)
<223>	AtMSH6 full-length cDNA and deduced sequence of the encoded
	polypeptide
400	20
<400>	30
aaaagttgag	ccctgaggag tatcgtttcc gccatttcta cgacgcaagg cgaaaatttt 60
tggcgccaat	ctttccccc tttcgaattc tctcagctca aaacatcgtt tctctctcac 120
tctctctcac	aattccaaaa a atg cag cgc cag aga tcg att ttg tct ttc 171 Met Gln Arg Gln Arg Ser Ile Leu Ser Phe

ttc Phe	caa Gln	aaa Lys	ccc Pro	acc Thr 15	gcg Ala	gcg Ala	act Thr	acg Thr	aag Lys 20	ggt Gly	ttg Leu	gtt Val	tcc Ser	ggc Gly 25	gat Asp	219
gct Ala	gct Ala	agc Ser	ggc Gly 30	Gly	ggc Gly	ggc Gly	agc Ser	gga Gly 35	gga Gly	cca Pro	cga Arg	ttt Phe	aat Asn 40	gtg Val	aag Arg	267
gaa Glu	GJÀ aaa	gat Asp 45	gct Ala	aaa Lys	ggc Gly	gac Asp	gct Ala 50	tct Ser	gta Val	cgt Arg	ttt Phe	gct Ala 55	gtt Val	tcg Ser	aaa Lys	315
tct Ser	gtc Val 60	gat Asp	gag Glu	gtt Val	aga Arg	gga Gly 65	acg Thr	gat Asp	act Thr	cca Pro	ccg Pro 70	gag Glu	aag Lys	gtt Val	ccg Pro	363
cgt Arg 75	cgt Arg	gtc. Val	ctg Leu	ccg Pro	tct Ser 80	gga Gly	ttt Phe	aag Lys	ccg Pro	gct Ala 85	gaa Glu	tcc Ser	gcc Ala	gst Gly	gat Asp 90	411
gct Ala	tcg Ser	tcc Ser	ctg Leu	ttc Phe 95	tcc Ser	aat Asn	att	atg Met	cat His 100	aag Lys	ttt Phe	gta Val	aaa Lys	gtc Val 105	gat Asp	459
gat Asp	cga Arg	gat Asp	tgt Cys 110	tct Ser	gga Gly	gag Glu	agg Arg	agc Ser 115	cga Arg	gaa Glu	gat Asp	gtt Val	gtt Val 120	ccg Pro	ctg Leu	507
aat Asn	gat Asp	tca Ser 125	tct Ser	cta Leu	tgt Cys	atg Met	aag Lys 130	gct Ala	aat Asn	gat Asp	gtt Val	att Ile 135	cct Pro	caa Gln	ttt Phe	555
cgt Arg	ser	aat Asn	aat Asn	ggt Gly	aaa Lys	act Thr 145	caa Gln	gaa Glu	aga Arg	aac Asn	cat His 150	gct Ala	ttt Phe	agt Ser	ttc Phe	603
agt Ser 155	GJA aaa	aga Arg	gct Ala	gaa Glu	ctt Leu 160	aga Arg	tca Ser	gta Val	gaa Glu	gat Asp 165	ata Ile	gga Gly	gta Val	gat Asp	ggc Gly 170	651
gat Asp	gtt Val	cct Pro	ggt	cca Pro 175	gaa Glu	aca Thr	cca Pro	GJA aaa	atg Met 180	cgt Arg	cca Pro	cgt Arg	gct Ala	tct Ser 185	cgc Arg	699
ttg Leu	aag Lys	cga Arg	gtt Val 190	ctg Leu	gag Glu	gat Asp	gaa Glu	atg Met 195	act Thr	ttt Phe	aag Lys	gag Glu	gat Asp 200	aag Lys	gtt Val	747
			gac Asp													795

Cys								gaa Glu									843
gag Glu 235			-			-	_		_	_	_						891
								ata Ile									939
								tgg Trp 275									987
								ej aaa									1035
								gag Glu									1083
agt Ser 315	ggt Gly	gtg Val	gga Gly	aaa Lys	tgc Cys 320	aga Arg	cag Gln	gtt Val	ggt	atc Ile 325	tct Ser	gaa Glu	agt Ser	Gly	ata Ile 330	ş.	1131
										Gly					cga Arg		1179
								caa Gln 355	Ala					Ala	aat Asn		1227
			Pro					Gln					Ser		gca Ala		1275
agc Ser	gag Glu 380	Gly	aac Asn	atc Ile	Gly	cct Pro 385	Asp	gcc Ala	gtc Val	cat His	Ctt Leu 390	Leu	gct Ala	ata Ile	aaa Lys		1323
g ag Glu 395	Ile	aaa Lys	atg Met	gag Glu	cta Leu 400	Gln	aag Lys	tgt Cys	tca Ser	act Thr 405	Val	tat Tyr	gga Gly	ttt Phe	gct Ala 410		1371
					Ala					Val					gat Asp		1419

					gct Ala											1467
					gac Asp											1515
					acg Thr											1563
					atg Met 480											1611
					gga Gly							_				1659
_	_	-	_		cta Leu		-	•	_	-	_		_	_		1707
					cat His			-		-		_	_	_		1755
					ttt Phe				-			-				1803
	-		_	_	atg Met 560	_								-	_	1851
					GJA aaa											1899
					cga 'Arg											1947
					atc Ile											1995
					agt Ser							Tyr				2043

					aga Arg 640											2091
tca Ser	tca Ser	gcc Ala	tct Ser	gtg Val 655	ttg Leu	cct Pro	gct Ala	ctt Leu	ctg Leu 660	Gly aaa	aaa Lys	aaa Lys	gtg Val	ctg Leu 665	aaa Lys	2139
					ttt Phe											2187
	-	_	_	_	gct Ala		_	_	_			_	_	_	_	2235
				_	aaa Lys ·					_			_			2283
					caa Gln 720											2331
		_			gat Asp			_	-		-	_				2379
			-		ttt Phe		-	_	_						_	2427
				-	tgc Cys		-	_	_	_			-		_	2475
_	_			_	gga Gly	_	_	_			_				gaa Glu	2523
		-		-	cag Gln 800		_									2571
					cat His					Ala						2619
					ata Ile									Ser		2667

agc Ser	att Ile	cat His 845	cct Pro	cgg Arg	tca Ser	ttg Leu	tta Leu 950	ctg Leu	acg Thr	gga Gly	cca Pro	aac Asn 855	atg Met	ggc	gga Gly	2715
aaa Lys	cca Ser 860	act Thr	ctt Leu	ctt Leu	cgt Arg	gca Ala 865	aca Thr	tgt Cys	ctg Leu	gcc Ala	gtt Val 870	atc Ile	ttt Phe	gcc Ala	caa Gln	2763
ctt Leu 875	ggc	tgc Cys	tac Tyr	gtg Val	ccg Pro 880	tgt Cys	gag Glu	tct Ser	tgc Cys	gaa Glu 885	Ile	tcc Ser	ctc Leu	gtg Val	gat Asp 890	2811
act Thr	atc Ile	ttc Phe	aca Thr	agg Arg 895	ctt Leu	ggc Gly	gca Ala	tct Ser	gat Asp 900	aga Arg	acc Ile	atg Met	aca Thr	gga Gly 905	gag Glu	2859
agt Ser	acc Thr	ttt Phe	ttg Leu 910	gta Val	gaa Glu	tgc Cys	act Thr	gag Glu 915	aca Thr	gcg Ala	tca Ser	gtt Val	ctt Leu 920	cag Gln	aat Asn	2907
gca Ala	act Thr	cag Gln 925	gat Asp	tca Ser	cta Leu	gta Val	atc Ile 930	ctt Leu	gac Asp	gaa Glu	ctg Leu	ggc Gly 935	aga Arg	gga Gly	act Thr	2955
agt Ser	act Thr 940	ttc Phe	gat Asp	gga Gly	tac Tyr	gcc Ala 945	acc Ile	gca Ala	tac Tyr	tcg Ser	gtt Val 950	ttt Phe	cgt Arg	cac His	ctg Leu	3003
gta Val 955	gag Glu	aaa Lys	gtt Val	caa Gln	tgt Cys 960	cgg Arg	atg Met	ctc Leu	ttt Phe	gca Ala 965	aca Thr	cat His	tac Tyr	cac	cct Pro 970	3051
ctc Leu	acc Thr	aag Lys	gaa Glu	ttc Phe 975	gcg Ala	tct Ser	cac His	cca Pro	cgt Arg 980	gtc Val	acc Thr	ccg Ser	aaa Lys	cac His 985	atg Met	3099
gct Ala	tgc Cys	gca Ala	ttc Phe 990	aaa Lys	tca Ser	aga Arg	tct Ser	gat Asp 995	tat Tyr	caa Gln	cca Pro	Arg	ggt Gly L000	tgt Cys	gat Asp	3147
caa Gln	Asp.	cta Leu 1005	gtg Val	ttc Phe	ttg Leu	Tyr	cgt Arg 010	tta Leu	acc Thr	gag Glu	Gly	gct Ala 1015	tgt Cys	cct Pro	gag Glu	3195
Ser	tac Tyr 1020	gga Gly	ctt Leu	caa Gln	Val	gca Ala L025	ctc Leu	atg Met	gct Ala	Gly	ata Ile 1030	cca Pro	aac Asn	caa Gln	gtg Val	3243
gtt Val 1039	Glu	aca Thr	gca Ala	Ser	ggt Gly L040	gct Ala	gct Ala	caa Gln	Ala	atg Met 1045	aag Lys	aga Arg	tca Ser	Ile	CJA GJA GGG	3291

gga aac ttc Glu Asn Phe	aag tca agt gag cta Lys Ser Ser Glu Leu 1055	aga tot gag Arg Ser Glu 1060	ttc tca agt ctg cat Phe Ser Ser Leu His 1065	3339
Glu Asp Trp	ctc aag tca ttg gtg Leu Lys Ser Leu Val 1070	ggt att tct Gly Ile Ser 1075	cga gtc gcc cac aac Arg Val Ala His Asn 1080	3387
aat gcc ccc Asn Ala Pro 1085	att ggc gaa gat gac Ile Gly Glu Asp Asp 1090	Tyr Asp Thr	ttg ttt tgc tta tgg Leu Phe Cys Leu Trp 1095	3435
	aaa too tot tac tgt Lys Ser Ser Tyr Cys 1105		taaatggcta	3478
tgacataaca	ctatctgaag ctcgttaag	t cttttgcctc	tctgatgttt attcctctta	3538
aaaaatgctt	atatatcaaa aaattgttt	c ctcgattaaa	aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	3598
aaaaaaa				360 %
			•	
<210>	31			
<211>	1109			4/,
<212>	PRT			
<213>	Arabidopsis thalia	<i>na</i> ecotype Co	olumbia	
<223>	Polypeptide MSH6			
<400>	31			
Met Gln Arg	Gln Arg Ser Ile Leu 5	Ser Phe Phe	Gln Lys Pro Thr Ala 15	
Ala Thr Thr	Lys Gly Leu Val Ser	Gly Asp Ala 25	Ala Ser Gly Gly Gly 30	
Gly Ser Gly 35	Gly Pro Arg Phe Asn 40	Val Arg Glu	Gly Asp Ala Lys Gly 45	
Asp Ala Ser 50	Val Arg Phe Ala Val 55	Ser Lys Ser	Val Asp Glu Val Arg	
Gly Thr Asp 65	Thr Pro Pro Glu Lys	Val Pro Arg 75	Arg Val Leu Pro Ser 80	
Gly Phe Lys	Pro Ala Glu Ser Ala 85	Gly Asp Ala 90	Ser Ser Leu Phe Ser 95	
Asn Ile Met	His Lys Phe Val Lys 100	Val Asp Asp 105	Arg Asp Cys Ser Gly 110	
Glu Arg Ser 115	Arg Glu Asp Val Val 120	Pro Leu Asn	Asp Ser Ser Leu Cys 125	

Mec	Lys 130	Ala	Asn	Asp	Val	Ile 135	Pro	Gln	Phe	Arg	Ser 140	Asn	Asn	Gly	Lys
Thr 145	Gln	Glu	Arg	Asn	His 150	Ala	Phe	Ser	Phe	Ser 155	Gly	Arg	Ala	Glu	Leu 160
Arg	Ser	Val	Glu	Asp 165	Ile	Gly	Val	Asp	Gly 170	Asp	Val	Pro	Gly	Pro 175	Glü
Thr	Pro	Gly	Met 180	Arg	Pro	Arg	Ala	Ser 185	Arg	Leu	Lys	Arg	Val 190	Leu	Glu
Asp	Glu	Met 195	Thr	Phe	Lys	Glu	Asp 200	Lys	Val	Pro	Val	Leu 205	Asp	Ser	Asn
Lys	Arg 210	Leu	Lys	Met	Leu	Gln 215	Asp	Pro	Val	Cys	Gly 220	Glu	Lys	Lys	Glu
Val 225	Asn	Glu	Gly	Thr	Lys 230	Phe	Glu	Trp	Leu	Glu 235	Ser	Ser	Arg	Ile	Arg 240
Asp	Ala	Asn	Arg	Arg 245	Arg	Pro	Asp	Asp	Pro 250	Leu	Tyr	Asp	Arg	Lys 255	Thr
Leu	His	Ile	Pro 260	Pro	Asp	Val	Phe	Lys 265	Lys	Met	Ser	Ala	Ser 270	Gln	Lys
Gln	Tyr	Trp 275	Ser	Val	Lys	Ser	Glu 280	Tyr	Met	Asp	Ile	Val 285	Leu	Phe	Phe
Lys	Val 290	Gly	Lys	Phe	Tyr	Glu 295	Leu	Tyr	Glu	Leu	Asp 300	Ala	Glu	Leu	Gly
His 305	Lys	Glu	Leu	_	Trp 310	Lys	Met	Thr	Met	Ser 315	Gly	Val	Gly	Lys	Cys 320
Arg	Gln	Val	Gly	Ile 325	Ser	Glu	Ser	_	Ile 330	_	Glu	Ala	Val	Gln 335	-
Leu	Leu	Ala	Arg 340	Gly	Tyr	Lys	Val	Gly 345	Arg	Ile	Glu	Gln	Leu 350	Glu	Thr
Ser	Asp	Gln 355	Ala	Lys	Ala	Arg	Gly 360	Ala	Asn	Thr	Ile	Ile 365	Pro	Arg	Lys
Leu	Val 370	Gln	Val	Leu	Thr	Pro 375	Ser	Thr	Ala	Ser	Glu 380	Gly	Asn	Ile	Gly
Pro 385	Asp	Ala	Val	His	Leu 390	Leu	Ala	Ile	Lys	Glu 395	Ile	Lys	Met	Glu	Leu 400
Gln	Lys	Cys	Ser	Thr		Tyr	Gly				Val	_	Cys	Ala 415	

Leu	Arg	Phe	Trp 420	Val	Gly	Ser	Ile	Ser 425		Asp	Ala	Ser	Cys 430	Ala	Ala
Leu	Gly	Ala 435	Leu	Leu	Met	Gln	Val 440	Ser	Pro	Lys	Glu	Val 445	Leu	туг	Asp
Ser	Lys 450	Gly	Leu	Ser	Arg	Glu 455	Ala	Gln	Lys	Ala	Leu 460	Arg	Lys	Tyr	Thr
Leu 465	Thr	Gly	Ser	Thr	Ala 470	Val	Gln	Leu	Ala	Pro 475	Val	Pro	Gln	Val	Met 480
Gly	Asp	Thr	Asp	Ala 485	Ala	Gly	Val	Arg	Asn 490		Ile	Glu	Ser	Asn 495	Gly
Tyr	Phe	Lys	Gly 500	Ser	Ser	Glu	Ser	Trp 505	Asn	Cys	Ala	Val	Asp 510	Gly	Leu
Asn	Glu	Cys 515	Asp	Val	Ala	Leu	Ser 520	Ala	Leu	Gly	Glu	Leu 525	Ile	Asn	His
Leu	Ser 530	Arg	Leu	Lys	Leu	Glu 535	Asp	Val	Leu	Lys	His 540	Gly	Asp	Ile	Phe
Pro 545	Tyr	Gln	Val	туг	Arg 550	Gly	Cys	Leu	Arg	Ile 555	Asp	Gly	Gln	Thr	Met 560
Val	Asn	Leu	Glu	Ile 565	Phe	Asn	Asn	Ser	Cys 570	Asp	Gly	Gly	Pro	Ser 575	Gly
Thr	Leu	Tyr	Lys 580	Tyr	Leu	Asp	Asn	Cys 585	Val	Ser	Pro	Thr	Gly 590	Lys	Arg
Leu	Leu	Arg 595	Asn	Trp	Ile	Cys	His 600	Pro	Leu	Lys	Asp	Val 605	Glu	Ser	Ile
Asn	Lys 610	Arg	Leu	Asp	Val	Val 615	Glu	Glu	Phe	Thr	Ala 620	Asn	Ser	Glu	Ser
Met 625	Gln	Ile	Thr	Gly	Gln 630	Tyr	Leu	His	Lys	Leu 635	Pro	Asp	Leu	Glu	Arg 640
Leu	Leu	Gly	Arg	Ile 645	Lys	Ser	Ser	Val	Arg 650	Ser	Ser	Ala	Ser	Val 655	Leu
Pro	Ala	Leu	Leu 660	Gly	Lys	Lys	Val	Leu 665	Lys	Gln	Arg	Val	Lys 670	Ala	Phe
Gly	Gln	Ile 675	Val	Lys	Gly	Phe	Arg 680	Ser /			Asp	Leu 685	Leu	Leu	Ala
Leu	Gln 690	Lys	Glu	Ser	Asn	Met 695	Met	Ser	Leu	Leu	Tyr 700	Lys	Leu	Cys	Lys

105	Pro	Ile	Leu	Val	Gly 710	Lys	Ser	Gly	Leu	Glu 715	Leu	Phe	Leu	Ser	Glr 720
Phe	Glu	Ala	Ala	Ile 725	Asp	Ser	Asp	Phe	Pro 730	Asn	Tyr	Gln	Asn	Gln 735	Ası
Val	Thr	Asp	Glu 740	Asn	Ala	Glu	Thr	Leu 745	Thr	Ile	Leu	Ile	Glu 750	Leu	Phe
Ile	Glu	Arg 755	Ala	Thr	Gln	Trp	Ser 760	Glu	Val	Ile	His	Thr 765	Ile	Ser	Cys
Leu	Asp 770	Val	Leu	Arg	Ser	Phe 775	Ala	Ile	Ala	Ala	Ser 780	Leu	Ser	Ala	Gly
Ser 785	Met	Ala	Arg	Pro	Val 790	Ile	Phe	Pro	Glu	Ser 795	Glu	Ala	Thr	Asp	Glr 800
Asn	Gln	Lys	Thr	Lys 805	Gly	Pro	Ile	Leu	Lys 810	Ile	Gln	Gly	Leu	Trp 815	His
Pro	Phe	Ala	Val 820	Ala	Ala	qsA	Gly	Gln 825	Leu	Pro	Val	Pro	Asn 830	Asp	Ile
		835	Glu				840					845			
Leu	Leu 850	Leu	Thr	Gly	Pro	Asn 855	Met	Ġly	Gly	Lys	Ser 860	Thr	Leu	Leu	Arg
865			Leu		870		٠			875					880
			Cys	885.					890					895	
Gly	Ala	Ser	Asp 900	Arg	Ile	Met	Thr	Gly 905	Glu	Ser	Thr	Phe	Leu 910	Val	Glu
Cys	Thr	Glu 915	Thr	Ala	Ser	Val	Leu 920	Gln	Asn	Ala	Thr	Gln 925	Asp	Ser	Leu
Val	1le 930	Leu	Asp	Glu	Leu	Gly 935	Arg	Gly	Thr	Ser	Thr 940	Phe	Asp	Gly	Tyr
Ala 945	Ile	Ala	Tyr	Ser	Val 950	Phe	Arg	His	Leu	Val 955	Glu	Lys	Val	Gln	Cys 960
Arg	Met	Leu	Phe	Ala 965	Thr	His	Tyr	His	Pro 970	Leu	Thr	Lys	Glu	Phe 975	Ala
Ser	His	Pro	Arg 980	Val	Thr	Ser	Lys	His 985	Met	Ala	Суѕ	Ala	Phe 990	Lys	Ser

Arg Ser Asp T 995	Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu 1000 1005	
Tyr Arg Leu T 1010	Thr Glu Gly Ala Cys Pro Glu Ser Tyr Gly Leu Gln Val 1015 1020	
Ala Leu Met A 1025	Ala Gly Ile Pro Asn Gln Val Val Glu Thr Ala Ser Gly 1030 1035 1040	
Ala Ala Gln A	Ala Met Lys Arg Ser Ile Gly Glu Asn Phe Lys Ser Ser 1045 1050 1055	
_	Ser Glu Phe Ser Ser Leu His Glu Asp Trp Leu Lys Ser 1060 1065 1070	
Leu Val Gly I 1075	Ile Ser Arg Val Ala His Asn Asn Ala Pro Ile Gly Glu 1080 1085	
Asp Asp Tyr A	Asp Thr Leu Phe Cys Leu Trp His Glu Ile Lys Ser Ser 1095 1100	•
Tyr Cys Val 9	Pro Lys	
	•	
<210>	32	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATHGENEA microsatellite	
<400>	32	
accatgcata go	cttaaactt cttg	24
<210>	33	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of ATHGENEA microsatellite	
<400>	33	
acataaccac a	aataggggt gc .	22

```
<210>
              34
 <211>
              18
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              Forward primer DMCIN-A for PCR on genomic DNA of Arabidopsis
 <223>
              thaliana ssp. Landsberg erecta "Ler"
 <400>
              34
 gaagcgatat tgttcgtg
                                                                           18
 <210>
              35
 <211>
              18
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
 <223>
              Reverse primer DMCIN-B for PCR on genomic DNA of Arabidopsis
              thaliana ssp. Landsberg erecta "Ler"
<400>
              35
agattgcgag aacattcc
                                                                           18
<210>
              36
<211>
              31
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Forward primer DMCIN-1 for PCR on genomic DNA of Arabidopsis
              thaliana ssp. Landsberg erecta "Ler"
<400>
             36
acgcgtcgac tcagctatga gattactcgt g
                                                                          31
<210>
             37
<211>
             29
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Reverse primer DMCIN-2 for PCR on genomic DNA of Arabidopsis
             thaliana ssp. Landsberg erecta "Ler"
<400>
             37
gctctagatt tctcgctcta agactctct
                                                                          29
```

<210>	38
<211>	32
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	38
gctctagagc	ttctcttaag taagtgattg at 32
<210>	39
<211>	48
<212>	DNA
<213>	Artificial sequence
<220>	•
<223>	Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	39
tcccccgggc	togagagate tecatggttt etteagetet atgaatee 48
<210>	. 40
<211>	26
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMCla for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	40
acgcgtcgac	gaattcgcaa gtgggg 26
<210>	41
<211>	38
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer DMClb for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	41

tccatggaga	terecegggt accgattrge tregaggg	
<210>	42	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATEAT1 SSLP marker in Arabidopsis thaliana subspecies	1
<400>	42	
gccactgcgt	gaatgatatg 20	
<210>	43	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of ATEAT1 SSLP marker in Arabidopsis thaliana subspecies	į.
<400>	43	
cgaacagcca	acattaattc cc 22	
<210>	44	
<211>	18	
<212>	DNA .	
<213>		
(213)	Artificial sequence	
<220>	·	
<223>	Forward primer for PCR amplification of NGA63 SSLP marker in Arabidopsis thaliana subspecies	
<400>	44	
aaccaaggca	cagaagcg 18	
<210>	45	
<211>	18	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA63 SSLP marker in	
	Arabidopsis thaliana subspecies	

<400>	45	
acccaagtga	togocaco	18
<210>	46	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA248 SSLP marke Arabidopsis thaliana subspecies	r in
<400>	46	
taccgaacca	aaacacaaag g	21
<210>	47	
<211>	22	•
<212>	DNA	
<213>	Artificial sequence	
<220>		•
<223>	Reverse primer for PCR amplification of NGA248 SSLP marke Arabidopsis thaliana subspecies	r in
<400>	47	
tctgtatctc	ggtgaattet ee	22
<210>	48	
<211> <212>	22	
<212> <213>	DNA -	
(213)	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA128 SSLP marke Arabidopsis thaliana subspecies	r in
<400>	48	
ggtctgttga	tgtcgtaagt cg	22
<210>	49	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		

	37	_
<223>	Reverse primer for PCR amplification of NGA128 SSLP marker in Arabidopsis thaliana subspecies	1
<400>	49	
atcttgaaa	c ctttagggag gg	
<210>	50	
<211>	22	
<212> <213>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies	
<400>	50	
ctgatctcac	ggacaatagt gc £22	
<210>	51	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies	
<400>	51	
ggctccataa	aaagtgcacc 20	
<210>	52	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA111 SSLP marker in Arabidopsis thaliana subspecies	
<400>	52	
ctccagttgg	aagctaaagg g	
<210>	53	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	

<220>		
<223>	Reverse primer for PCR amplification of NGA111 SSLP marke Arabidopsis thaliana subspecies	r in
< 400 >	53	
tgttttttag	gacaaatggc g	21
<210>	54	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA168 SSLP marke Arabidopsis thaliana subspecies	r in
<400>	54	
ccttcacatc	caaaacccac	20
<210>	55	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA168 SSLP marke Arabidopsis thaliana subspecies	: r in
<400>	55	
gcacataccc	acaaccagaa	20
<210>	56	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>-	Forward primer for PCR amplification of NGA1126 SSLP marke	٠,
	in Arabidopsis thaliana subspecies	- -
<400>	56	
cgctacgctt	ttcggtaaag	20

Carrier of the State

<210>	57
<211>	20
<212>	DNA
<213>	
	Artificial sequence
<220>	
<223>	Payone mi
1237	Reverse primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis thaliana subspecies
	in Arabidopsis thaliana subspecies
<400>	
(400)	57
gcacagte	ca agtcacaacc
Judagee	20
	20
<210>	58
<211>	20
<212>	
<213>	DNA
\&LJ >	Artificial sequence
<220>	
<223>	Plane : 1
(223)	Forward primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	58
aaagagato	ga gaatttggac
	20
<210>	59
<211>	23
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Reverse primar for pop
	Reverse primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	59
acatatcaas	atattaaagt agc
	23
<210>	60
<211>	18
<212>	DNA
<213>	
	Artificial sequence
<220>	
<223>	Forward primary
	Forward primer for PCR amplification of NGA168 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	
	60

tegtetactg	cactgeeg 18
<210>	61
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	61
gaggacatgt	ataggagcct cg . 22
<210>	62
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	62
tgacctcctc	ttccatggag 20
<210>	63
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	63
ttaacagaaa	cccaaagctt tc 22
<210×	
<210><211>	64
	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of AthUBIQUE SSLP marker in Arabidopsis thaliana subspecies

```
<400>
                64
  aggcaaatgt ccatttcatt g
                                                                             21
  <210>
               65
  <211>
               20
  <212>
               DNA
  <213>
               Artificial sequence
  <220>
               Reverse primer for PCR amplification of AthUBIQUE SSLP marker
  <223>
               in Arabidopsis thaliana subspecies
  <400>
               65
 acgacatggc agatttctcc
                                                                            20
 <210>
               66
 <211>
               21
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              Forward primer for PCR amplification of NGA172 SSLP marker in
 <223>
              Arabidopsis thaliana subspecies
 <400>
              66
 agetgettee ttatagegte e
                                                                            21
 <210>
              67
 <211>
              19
 <212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Reverse primer for PCR amplification of NGA172 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             67
catccgaatg ccattgttc
                                                                           19
<210>
             68
<211>
             21
<212>
             DNA
<213>
             Artificial sequence
<220>
```

<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA126	SSLP	marker in
-: 400 >	63		•		
gaaaaaacgc	tactttcgtg g			•	21
<210>	69				
<211>	22				
<212>	DNA				
<213>	Artificial sequence				
<220>					
<223>	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA126	SSLP	marker in
<400>	69				
caagagcaat	atcaagagca gc				22
<210>	70				•
<211>	20				
<212>	DNA				
<213>	Artificial sequence				
<220>					•
<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA162	SSLP	marker in
<400>	70				
catgcaattt	gcatctgagg				20
<210>	71 .				
<211>	22		•		
<212>	DNA				
<213>	Artificial sequence				
<220>					
<223>	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA162	SSLP	marker in
<400>	71				
ctctgtcact	cttttcctct gg				22
010					
<210>	72				•
<211>	21	•			
<212>	DNA				

الوقي المراكب عوالي

```
<220>
   <223>
                Forward primer for PCR amplification of NGA6 SSLP marker in
                Arabidopsis thaliana subspecies
   <400>
                72
   tggatttctt cctctcttca c
                                                                             21
  <210>
                73
  <211>
                21
  <212>
                DNA
  <213>
                Artificial sequence
  <220>
  <223>
               Reverse primer for PCR amplification of NGA6 SSLP marker in
               Arabidopsis thaliana subspecies
  <400>
               73
 atggagaage tracactgat c
                                                                            21
 <210>
               74
 <211>
               20
 <212>
               DNA
 <213>
               Artificial sequence
 <220>
 <223>
              Forward primer for PCR amplification of NGA12 SSLP marker in
              Arabidopsis thaliana subspecies
 <400>
              74
aatgttgtcc tcccctcctc
                                                                           20
<210>
              75
<211>
              22
<212>
              DNA
<213>
             Artificial sequence
<220>
<223>
             Reverse primer for PCR amplification of NGA12 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             75
tgatgctctc tgaaacaaga gc
```

<210>	76
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA8 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	76
gagggcaaat	ctttatttcg g
<210>	77
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	77
tggctttcgt	ttataaacat cc 22
<210>	78
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1107 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	78
gcgaaaaaac	aaaaaaatcc a
<210>	79
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
<400>	79

```
cgacgaatcg acagaattag g
                                                                             21
   <210>
                80
   <211>
                21
   <212>
                DNA
   <213>
               Artificial sequence
  <220>
  <223>
               Forward primer for PCR amplification of NGA225 SSLP marker in
               Arabidopsis thaliana subspecies
  <400>
               80
  gaaatccaaa tcccagagag g
                                                                            21
  <210>
               81
  <211>
               22
  <212>
               DNA
  <213>
               Artificial sequence
 <220>
 <223>
              Reverse primer for PCR amplification of NGA225 SSLP marker in
              Arabidopsis thaliana subspecies
 <400>
              81
 tetececact agettegege ce
                                                                           22
 <210>
              82
 <211>
              19
 <212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Forward primer for PCR amplification of NGA249 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             82
taccgtcaat ttcatcgcc
                                                                          19
<210>
             83
<211>
             22
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
            Reverse primer for PCR amplification of NGA249 SSLP marker in
            Arabidopsis thaliana subspecies
```

46

<400>	83
ggatocotaa	ctgtaaaatc cc 22
<210>	84
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	\cdot
<223>	Forward primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	84
aatcccagta	accaaacaca ca 22
<210>	85
<211>	20
<212>	DNA
<213>	Artificial sequence
	crricial seddence
<220>	
<223>	Reverse primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	85
cccagtctaa	ccacgaccac 20
<210>	86 ·
<211>	20
<211>	DNA
<213>	
(213)	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA151 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	86
gttttgggaa	gttttgctgg
<210>	87
<211>	24
<212>	DNA
<213>	Artificial sequence
<220>	·•
	···

e esta a esta e

```
Reverse primer for PCR amplification of NGA151 SSLP marker in
  <223>
               Arabidopsis thaliana subspecies
  <400>
               87
  cagtictaaaa gigagagtat gatg
                                                                            24
  <210>
               88
 <211>
               22
 <212>
               DNA
 <213>
               Artificial sequence
 <220>
              Forward primer for PCR amplification of NGA106 SSLP marker in
 <223>
              Arabidopsis thaliana subspecies
 <400>
              88
 gttatggagt ttctagggca cg
                                                                           22
 <210>
              89
 <211>
              20
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
 <223>
              Reverse primer for PCR amplification of NGA106 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              89
tgccccattt tgttcttctc
                                                                          20
<210>
              90
<211>
              20.
<212>
             DNA
<213>
           Artificial sequence
<220>
             Forward primer for PCR amplification of NGA139 SSLP marker in
<223>
             Arabidopsis thaliana subspecies
<400>
             90
agagctacca gatccgatgg
                                                                         20
<210>
             91
<211>
             21
<212>
             DNA
<213>
          F Artificial sequence
```

48

<220>	\cdot
<223>	Reverse primer for PCR amplification of NGA139 SSLP marker in Arabidopsis thaliana subspecies
< 400 >	91
ggtttcgttt	cactatccag g 21
<210>	92
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA76 SSLP marker in Arabidopsis thaliana subspecies
<400>	92
ggagaaaatg	tcactctcca cc 22
<210>	93
<211><212>	20
<212>	DNA
(213)	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA76 SSLP marker in Arabidopsis thaliana subspecies
<400>	93
aggcatggga	gacatttacg 20
	•
<210>	94
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for DCB amplification of amuseus corp
	Forward primer for PCR amplification of ATHSO191 SSLP marker in Arabidopsis thaliana subspecies
<400>	94
ctccaccaat	Catgcaaatg : 20

<210>	95	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Perezo primer for per	
	Reverse primer for PCR amplification of ATHSO191 SSLP in Arabidopsis thaliana subspecies	marker
<400>	95	
tgatgttga	t ggagatggtc a	21
_		
<210>	96	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		. •
<223>	Forward primer for PCR amplification of NGA129 SSLP ma	
	Arabidopsis thaliana subspecies	arker in
<400>	96	
tcaggaggaa	a ctaaagtgag gg	22
<210>	97	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA129 SSLP ma Arabidopsis thaliana subspecies	rker in
<400>	97	
cacactgaag	atggtcttga gg	22
<210>	98	•
<211>	8062	
<212>	DNA	•
<213>		
1220	Arabidopsis thaliana ecotype Columbia	
<220>		
<223>	Genomic DNA sequence of AtMSH6	
<400>	97	
ttttttggtt	gctaacaata aaggtatacg gttttatgtc atcaatataa ctatatataa	60

aagaaatgaa	a agatatatat	tgtttttca	tttatcaaac	aaaacaacaa	gactttttt	120
ttacttttta	a cattggtcaa	l caaaatacaa	gataaacgac	atogettaat	catttcccaa	180
ttttacccct	: aagtttaaca	cctagaacct	tctccatctt	cgcaagcaca	gcctgattag	240
gaacagcttt	accattctca	tattcctgaa	ctacctgagt	cctctcattg	atctgtttcg	300
ccaaatccgc	ttgtgacato	ttcttctcca	atctcgcttt	ctgtatcatc	aacctcacct	360
ctgctttcac	acgatccatc	gccgcaggct	ctgtttcttc	ttccagette	ttcgtgttaa	420
tcaccggaac	cgccgtagat	ttcccctttt	tgttcgaacc	ggcatcgaat	ttcttaaccg	480
tttgaaccgc	gacaccgttt	ctcagagctg	cgttaaccgc	tttcggatcg	cgtaggtctt	540
ggctcttttg	ttttgatttg	tggagaacta	ctggttccca	gtcttgtgtt	actgctcctg	600
ggtatetget	cggcatcgtc	gatgaattga	gagaaaggaa	caacgcgaaa	attttattaa	660
tctgagttt	gaaattgaga	aacgatgaag	atgaagaatg	ttgttgagag	gattgtgata	720
tttatatata	cgaagattgg	tttctggaga	attcgatcat	ctttttctcc	attttcgtct	780
ctggaacgtt	cttagagatg	attgacgacg	tgtcattatc	tgatttgcag	ttaaccaatg	840
ctttttgggt	tggattcgtg	gtacaccata	ttatccgatt	tggctcaatg	gttttatata	900
aatttggttt	teggtteggt	tatgagttat	cattaaaatt	aagctaacca	aaaattttcg	960
taaaatttat	ttcggtttca	attcggatcc	cttacttcca	gaaccgaatt	attcgaaacc	1020
ggggttagcc	gaaccgaata	ccaatgcctg	attgactcgt	tggctagaaa	gatccaacgg	1080
tatacaataa	tagaacataa	atcggacggt	catcaaagcc	tcaaagagtg	aacagtcaac	1140
aaaaaaagtt	gagccctgag	gagtatcgtt	tccgccattt	ctacgacgca	aggcgaaaat	1200
ttttggcgcc	aatctttccc	ccctttcgaa	ttctctcagc	tcaaaacatc	gtttctctct	1260
cactetetet	cacaattcca	aaaaatgcag	cgccagagat	cgattttgtc	tttcttccaa	1320
aaacccacgg	cggcgactac	gaagggtttg	gtttccggcg	atgctgctag	cggcgggggc	1380
ggcagcggag	accacgattt	aatgtgaagg	aaggggatgc	taaaggcgac	gcttctgtac	1440
gttttgctgt	ttcgaaatct	gtcgatgagg	ttagaggaac	ggatactcca	ccggagaagg	1500
ttccgcgtcg	tgtcctgccg	tctggattta	agccggctga	atccgccggt	gatgcttcgt	1560
ccctgttctc	caatattatg	cataagtttg	taaaagtcga	tgatcgagat	tgttctggag	1620
agaggtacta	atcttcgatt	ctcttaattt	tgttatcttt	agctggaaga	agaagattcg	1680

	tgtaatttgt tgtattcgtt ggagagattc tgattactgc attggatcgt tgtttacaaa	174(
	ttttcaggag ccgagaagat gttgttccgc tgaatgattc atctctatgt atgaaggcta	1800
	atgatgitat teeteaatti egiteeaata atggiaaaae teaagaaaga aaccatgett	1860
	rtagtttcag tgggagaget gaacttagat cagtagaaga tataggagta gatggegatg	1920
	tteetggtee agaaacacca gggatgegte caegtgette tegettgaag egagttetgg	1980
	aggatgaaat gacttttaag gaggataagg ttcctgtatt ggactctaac aaaaggctga	2040
	aaatgeteea ggateeggtt tgtggagaga agaaagaagt aaaegaagga accaaatttg	2100
	aatggettga gtettetega atcagggatg ceaatagaag aegteetgat gateeeettt	
	acgatagaaa gaccttacac ataccacctg atgttttcaa gaaaatgtct gcatcacaaa	2160
	agcaatattg gagtgttaag agtgaatata tggacattgt gcttttcttt aaagtggtta	2220
	gtaactatta atctagtgtt caatccattt cctcaatgtg atttgttcac ttacatctgt	2280
	ttacgttatg ctcttctcag gggaaatttt atgagctgta tgagctagat gcggaattag	2340
	gtcacaagga gcttgactgg aagatgacca tgagtggtgt gggaaaatgc agacaggtaa	2400
	attagttgaa acaactggcc tgcttgaatt attgtgtcta taaattttga caccaccttt	2460
	tgtttcaggt tggtatctct gaaagtggga tagatagaa	2520
	tgtttcaggt tggtatctct gaaagtggga tagatgaggc agtgcaaaag ctattagctc	2580
	gtgggtaagg gaaccatcat actttatgga attcgtttac tgctacttcg gctaggattt	2640
	aagaaatgga aatcacttca agcatcatta gttaggatcc tgagaactca ggatgttttc	2700
1	ttattcgtta tataataagt cttttcatca aggagtaaca aacaaaactt gcacaatatt	2760
1	tgtgtgctca ctggcaaggc atatataccc agctaacctt tgctagttca ctgtagtaac	2820
ć	agttacggat aatatatgtt tacttgtatg tggtaccctc attttgtctc tcatggaggc	2880
t	ttcaagcet tgtgttgaaa etggatagtt acatatgett ecaacagaaa etageatgea	2940
9	gattcatatg ctttcctatt ctactaatta tgtattgaca cactcgttgt ttcttttgaa	3000
а	gatataaag ttggacgaat cgagcagcta gaaacatctg accaagcaaa agccagaggt	3060
g	ctaatactg taagttttct tggataggtc aaggagagtg ttgcagactg tttttgatca	3120
t	ttottttte tgtacattae ttteatgetg taattaaete aatggetatt etggtetgat	3180
t.	atcagataa ttccaaggaa gctagttcag gtattaactc catcaacagc aagcgaggga	•
a	acatogggo otgatgoogt coatottott gotataaaag aggtttgtta tttacttatt	3240
	" - 22 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 -	3300

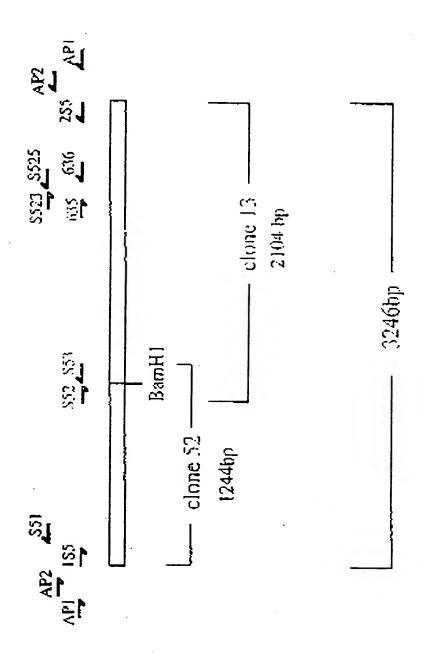
tatcttatca	cgttcagttc	atccaagtcc	tgaaaaatta	cactcttctt	taccaatctt	3360
ccatcaagct	gtgtaaagga	tttggaatta	gaaaatcatt	atttgatgct	ttgttttata	3420
tgcaagaggt	tcccttgaaa	agatctgttt	aagattcttt	gcacttgaaa	aattcaatct	3480
ttttaagtga	atcccctact	ttcttacaat	gatcatagtc	tgcaattgca	tgtcaagtaa	3540
tatcattcct	tgttactgca	tcccctctt	tcttaatgac	cattgtctat	gttgtgtttg	3600
tctcgtgtgc	tggagaaaat	gatagctgat	ccaagctgta	cattatcatg	attaagtagc	3660
tgctcaggaa	ttgcctttgg	ttacattgcc	taatggtttg	atgtcaattt	ttcttctgaa	3720
tctttattt	agatcaaaat	ggagctacaa	aagtgttcaa	ctgtgtatgg	atttgctttt	3780
gttgactgtg	ctgccttgag	gttttgggtt	gggtccatca	gcgatgatgc	atcatgtgct	3840
gctcttggag	cgttattgat	gcaggtaagc	aagtgtattc	tgtatcttat	gtgtaccatg	3900
tgacttcctg	tgcatatatt	tgggttgcag	gaactaattc	tgaatcacca	tttggtatgt	3960
tttttccagg	tttctccaaa	ggaagtgtta	tatgacagta	aaggtaaact	gcttgtätcg	4020
ccagttgttt	tgttaaacag	aatttaaggt	aaatgacact	ggttaattta	aagtgcatac	4080
atgttgaaat	attgcagggc	tatcaagaga	agcacaaaag	gctctaagga	aatatacgtt	4140
gacaggtacc	atttcagtag	gcaagctaac	tgacaattta	accgctcacc	gaatgatagg	4200
tctcttaaac	attgctaatg	tagatgatgt	ttatgtttca	atctaatagg	gtctacggcg	4260
gtacagttgg	ctccagtacc	acaagtaatg	ggggatacag	atgctgctgg	agttagaaat	4320
ataatagaat	ctaacggata	ctttaaaggt	tcttctgaat	catggaactg	tgctgttgat	4380
ggtctaaatg	aatgtgatgt	tgcccttagt	gctcttggag	agctaattaa	tcatctgtct	4440
aggctaaagg	tgtgttggct	tgtttagttt	ttgcttttca	caaattaagc	aaaggaactt	4500
ttcataactt	acagtttcta	tctacttgca	gctagaagat	gtacttaagc	atggggatat	4560
ttttccatac	caagtttaca	ggggttgtct	cagaattgat	ggccagacga	tggtaaatct	4620
tgagatattt	aacaatagct	gtgatggtgg	tccttcaggc	aagtgcatat	ttcttttttg	4680
ataacttcaa	ctagagggca	gacatagaag	gaaaaattct	aatacttcgt	acggatctcc	4740
agtaagtaat	agccgatttt	tgtttaccta	tgtagggacc	ttgtacaaat	atcttgataa	4800
ctgtgttagt	ccaactggta	agcgactctt	aaggaattgg	atctgccatc	cactcaaaga	4860
tgtagaaagc	atcaataaac	ggcttgatgt	agttgaagaa	ttcacggcaa	actcagaaag	4920

tatgeaaate actggecagt atetecaeaa aettecagae ttagaaagae tgeteggaeg	
catcaagtet agegttegat catcageete tgtgttgeet getettetgg ggaaaaaagt	
gctgaaacaa cgagtaagta tcaatcacaa gttttctgag taatgccttc catgagtagt	
ataggactaa aacattacgg gtctagctaa agactgttct ccttcttttg caatgtctgg	
ttattcatta catttctctt aacttattgc attgcaggtt aaagcatttg ggcaaattgt	
gaaagggttc agaagtggaa ttgatctgtt gttggctcta cagaaggaat caaatatgat	
gagtttgctt tataaactct gtaaacttcc tatattagta ggaaaaagcg ggctagagtt	5340
atttettet caattegaag cagecataga tagegaettt ecaaattate aggtgeeeat	5400
ctatctttca tactttacaa caaaatgtct gtcactactc aaagcaatgc atatggctta	5460
gateteaaet eacaccega ggateetaaa gggatttget tittatteet aatgittitg	5520
gatggtttga tttatttcta acttgaactt attaatcttg taccagaacc aagatgtgac	5580
agatgaaaac gctgaaactc tcacaatact tatcgaactt tttatcgaaa gagcaactca	5640
atggtctgag gtcattcaca ccataagctg cctagatgtc ctgagatctt ttgcaatcgc	5700
agcaagtete tetgetggaa geatggeeag geetgttatt ttteeegaat cagaagetae	5760
agatcagaat cagaaaacaa aagggccaat acttaaaatc caaggactat ggcatccatt	5820
tgcagttgca gccgatggtc aattgcctgt tccgaatgat atactccttg gcgaggctag	5880
aagaagcagt ggcagcattc atcctcggtc attgttactg acgggaccaa acatgggcgg	5940
aaaatcaact cttcttcgtg caacatgtct ggccgttatc tttgcccaag tttgtatact	6000
cgttagataa ttactctatt ctttgcaatc agttcttcaa catgaataat aaattctgtt	6060
ttctgtctgc agcttggctg ctacgtgccg tgtgagtctt gcgaaatctc cctcgtggat	6120
actatettea caaggettgg egeatetgat agaateatga eaggagagag taagttttgt	6180
totoaaaata coaattooto gaactattta otoagatttt gtotgattgg acaaggtggt	6240
tttgcttttt tttaggtacc tttttggtag aatgcactga gacagcgtca gttcttcaga	6300
atgcaactca ggattcacta gtaatccttg acgaactggg cagaggaact agtactttcg	63.60
atggatacge cattgeatae teggtaacet getettetee tteaaettat aettgttgat	6420
caacaaaaac atgcaattca ttttgctgaa acttattgat ttatatcagg tttttcgtca	6480
cctggtagag aaagttcaat gtcggatgct ctttgcaaca cattaccacc ctctcaccaa	6540

ggaattegeg	tctcacccac	gtgtcacctc	gaaacacatg	gcttgcgcat	tcaaatcaag	6600
acctgattat	caaccacgtg	gttgtgatca	agacctagtg	ttcttgtacc	gtttaaccga	6660
gggagcttgt	cctgagagct	acggacttca	agtggcactc	atggctggaa	taccaaacca	6720
agtggttgaa	acagcatcag	gtgctgctca	agccatgaag	agatcaattg	gggaaaactt	6780
caagtcaagt	gagctaagat	ctgagttctc	aagtctgcat	gaagactggc	tcaagtcatt	6840
ggtgggtatt	tctcgagtcg	cccacaacaa	tgcccccatt	ggcgaagatg	actacgacac	6900
tttgttttgc	ttatggcatg	agatcaaatc	ctcttactgt	gttcccaaat	aaatggctat	6960
gacataacac	tatctgaagc	tegttaagte	ttttgcttct	ctgatgttta	ttcctcttaa	7020
aaaatgctta	tatatcaaaa	aattgtttcc	tcgattataa	caagattata	tatgtatctg	7080
tcggtttagc	tatggtatat	aatatatgta	tgttcatgag	attggtcaag	agaaatactc	7140
acaaacagta	tattaagaag	gaaatatgtt	tatgcattaa	tttaagtttc	aagataaact	7200
gcaaataacc	tcgactaaag	ttgcaaagac	caaacacaaa	ttacaaaact	tataagactt	7260
aagttctgaa	ttccctaaaa	ccaaaaaaaa	aaacagaaca	tattttgttg	catctacaaa	7320
caacacaaac	ctacatagtt	tataacttac	tcatcactga	gattaacatc	agaatcattc	7380
tccatttctt	catcttcact	ctcatcatca	tcaccaccac	catgatgatt	CTCCTCCT	7440
tcacgtaacc	tagcaatctc	actctgagct	ctatcaacaa	tctgcttctt	ctgcaactcc	7500
aaatctctct	gaaaatcagc	tctcatcttc	tccaactcct	tcatttgctc	tttcttactc	7560
ttctccatct	tctcataaac	cttcccaaac	ctctcaacag	aatccgccaa	catcttatac	7620
gaagcagcgt	cattaacctt	cttcctctcg	tactcaacct	catcatcctc	atcctcctcc	7680
tcttcagaat	caccaggact	atccatcatc	tcatcaaacc	cattagactt	atctaaataa	7740
accttagtgt	tcataaacac	aaactcacct	gaatcaacac	cacaagctaa	acctaaatcc	7800
gacttgggcg	aaacacaaag	caacatatcc	aacttattga	aaaacgacca	tttactțgaa	7860
cctaaacctg	atttctcaac	cttaatcttc	tetttetat	acttcctctt	caagtcatca	7920
atcattctcc	tacattgcgt	ctcagatttc	tccatcctta	gctcctcact	cactttctca	7980
gctacttcat	tccaatcctc	gttcctcaaa	ctccttctac	ccaattgcaa	aaacctatct	8040
cccaaactt	caagcaacac	aa				8062

•		14.0				V. (g.
				* T		<i>:</i> .
				•		
				÷		
					•	
÷			•			
*		•				
Ji	.† ÷					
			,			
	· ·		e.			4

Figure 1



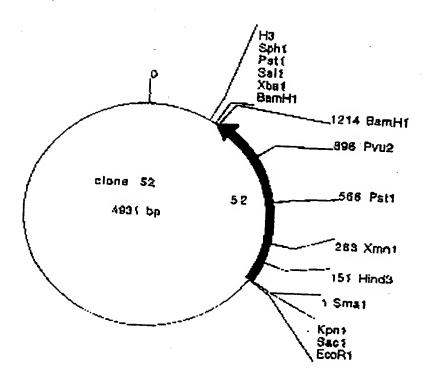


Figure 2

Commente/References: 52= 3' side of S5 (AIMSHa) 1244bp in pUC18/Sma1

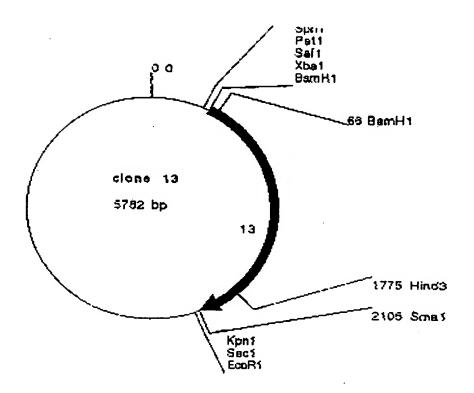


Figure 3

Commenta/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Smail

*
0
6
7

30	144 15	204 35	\$\$ \$\$ \$\$	324	38 4 95	4 1 1 1 5	504 335	564 155	624	684 195	744 215	804 235	864 255	924 275
GCGCGCGARARITGGCAACCCRAGITGGCCATGGCCACGACCACGACCTTGCATIICICIIAAACGGAGGA	22 2	933 ~	CTC	GAT O	GTT V	AAG K	455 6	TIC	ភ្ជ ១ ៩	ಕ್ಷನ್ನು ಕ	445 3	A.A.T N	GA.A.	GTT V
950	100 pt	ლ		22 4	TAT (CTA 7				GCT C				
THA	TIC	993 A	GRC C	GTA C	GPA Y	0.040 0.33						s AGT	r ATT	GTT V
il Ci							F TTC	ექე ლ		5 =				S. E.
TTE	r Frc	A ACA	S TCC	SCC A		OIG A	AGA	GAT D	ក្ស ស	TCC) ()	ដ្ឋា	76.7 C	GGT G
נטט	153 K	TCA	CTC	AHC N	S S S	ر د تر		ATG M	₹ 58 ¥	S X X	Ath K	631	ည္သ	A CA
2031	77.5	స్ట్రా	CTT 1	****	TCG	ر د د	95 e	اب م ت ش	150 W	ATT	7 L	11T	ATT.	300 y
HS S3	rat	5 3	zag K	ACT T	ည္သ	ම දුර ලද	7.5.7 X	GCT	S S S	600 A	Test 3	63.F 1	ACA T	ert ? I
CC	35C	50 50 50 50 50 50 50 50 50 50 50 50 50 5	R R	ည်း သူ့ အ	3.55 E	445	150	TAC (14 14 14 14 14 14 14 14 14 14 14 14 14 1	<u>3</u>	TTG T L	3 3 3 9	GAG A	Gra y
2025	S S O	61.8°	AAG (ccr (CTS C	TTG G	Grr G	ATT T I	TTC C					
MGCC	200	0 000 d	700 X	TCT C						P ACT	300 E	इ.स.च	S S	GTT V
3	i) o g				4 PYT	గ్రా త	ig m	689	ች ት	A.B.	rcg	GCT	argg X) (9)
ည်	A X	5 8.43°	, CCT		. e. C.A.	٠ ١ ١	हें इ	TTG i	TTC L	ACT T	CTG I.	် ပြော	GT 7:	GTT V
AGT.	3 0	000 9	i i	ž. Ž	S O	TAC Y	A DE E	GTG V	ಗ್ರವಿ ಕೃತ್ತಿ	80	ភពិភ ភ	16t C	AGA R	GTT V
533	* AAG	684 8	TIC	CC7	D T	A A A	17G L	ر دور	TTT E	A.B.G K	200 8	: : : :	GAG S	69
5	ပ္သ ဗ္	CAC FF	3	X X	T.T.	A D	GTT	ට් ජ	Į,	GTC	TIC	GGT	GAT	010 010
attg	atg M	ACT T	GTA V	A X	A CA	TCG	15 ≯	GC&	3	67.B	îti E	AGT S	STG >	a: 15 25 26
Graa	TT XA	9 5 8	ACT.	200	30	TC.	GRT	ATC	GTG	ה ה ה		ATA I	GTT (GTT A
ပွဲ ပွဲ	Tabagcaatt	3CC) 4	7C#	1 55	TCA S	200 *	ទ្ធខ្លួ	AGT	ATT .)	GA: 1	rgr c	GAT G
S S S	us ta	ž X	S	900	TTA L	₩CG	TAC	900 8	900	DE X	#CC () 945 3	GTT 1	TIT G
cctangraa c	grttrcgra	۵ په	ATA I	8	PAT N	45 m	AAG 7	CAC (ACG O	TEC P	CGC #	उटा व इ	TTG G L v	
KIN KIN	W.T.E.	PAR K	AAG J	000	700	222	AGC A	644 6 8		स्				B SGT
Ü	U		4 ×	Ŭ Æ	ប្ធ	ပိုစ	¥ W	შ დ	A TG	GGA G	A.P.C.	ව ≉	TTC	%TG ₩
-	\$ T	145 16	36	265 56	325 76	385 96	445	505 136	565 156	625 176	665 196	745 216	305	56

ଜ 984 295	G 1044 315	1104 335	9. 1164 355	T 1224 375	r 1234 395	1344	(T),404 435	B 1464 455	ж 1524 475	3C 1584 495	T 1644 515	m 1704 535	7 1764 555	575 575
rtg.	ក្ ភ	S SGC	TT T	h GTT V	car H	3 TCP.	t AAT N	4 GTB V	ATA S	r TCC S	J.L.L.	1867 1	s GCT	A AGG
₽GC S	1 16	₩ 1 1Ĉ	N N	ដូ ។	16c	TIG	rar K	ACA T	TTG 1	S	E B	150 K	i High	ATG M
TTG L	FII	161	GCT G	1 13 13 13 13 13 13 13 13 13 13 13 13 13	7 T	rc1 9	GTG V	CTT	AAT N	AGT	r P	క్ష చ	etg B	GMA E
at I	r. X	GRI	អ្នក	ညီ ဂိ	ACG 7	ဂ ဂ် ဂ	GTT V	F.	보 요 요	CAT *	a S	at'i I	५ १५	fcT S
GIG V	୍ତୁ ଓଡ଼	ट्रा र	ಕ್ಟರ ಕ	ICT S	(13)	T 7 T F	ဂ္ သ	SEC H	C#1	÷cī s	ទីវិនិ ទ	GAT ប	() ()	CAC D
50 T	¥. F-	a Line	D a H	A TG	ည ်း	S S	T F: J	AA.T	55°	ig o	att 1	TC! S	ATT I	5 a
្ន ទី ព	5	5 5 7 7 7 7 7	4.5.7 4.2.7	S S S	ដូន	13 4. 17 17	0 0 0	.9.7G	Ĕ.,	27G	00% •	TC?	11C	K.S.
TI3 L	£43	(5) E	हरू स्ट्रिस	**************************************	့ သည် (၁)	9	300	ent M	CCI	360	AGN %	age R	13 13 13 13 13 13 13 13 13 13 13 13 13 1	rta 1
4:55 5	TC3	វិទ្ធិ ភ	TST	44.0 13.0 13.0 13.0 13.0 13.0 13.0 13.0 13	నే చేసి	1	CJG L	Ç∓ ∓	r F	scr 3	មួ មួយ	ICT S	i S	ეენ მ
25. S	CIT L	GTS ;	17% L	GCT A	GII	TAC Y	ACT T	ንኅር ೯	i) i) i) i)	ret S	TCT	atg K)) (*	CII L
#G# 80	ದಿ	ر نان هر	Ş s	BCT A	5 2 ₽	CTT.	£3.	arr.	GTG V	ETT I	CCJ G	GCT A	ሊ ፡ች ጽ	CGG CT
atg A	9 3 0	cr. v	rt. I	GAG E	ិក ភ	15.50 1	Ω Ω Ω	100 100 100 100	ဗ္ဗ ဗု	ភ្ន ភ្នំ ធ	g E	ğ	SCT	CKG D
11C	و و) 6 z	GTT >	CTG L	CAT TAT	ာ (၅ ရ) နှင့်	ရှိ () ()	ည္သမ	ű z	S	ក្នុង ក	116	ACT T	ATT
ZAT Z	נייד	ភ្នំ	្ត ស្វាស	X X	5	स्रुध	CIC	ici S	2008 8	GTT	GRT	GTC V	CGG R	2 2
GAT	TTG L	T CC	GAT D	atg N	ATG M	TTT E	acr T	मू ज	CII	SCT.	11G	act S	CAT M	F.A.G
ART	CTG	CCT	GTA	្ត ម្តី ដ	PAC E	6 68	ATG M	र इ	CIT	C.B.T.	GAG E	TCC S	rit F	80
TTC	545 3	430 b	SCA A	X X	A TG	TTT	છે. છે	GGA G	A 80	CIT	AGT S	CIC L	ATC I	939 «
245	GCT A	GCT A		GRT								GTG ♥	as a	1.1
445 5	ECC P	CAT		GAT		3 3 *	OF Z	နှင့် စ		GCT	CIC	CTC	ACA T	TIP
TAT Y	TCA .	SCB S	AAT (5.05 H	CP.T	CIC	AGT	AAT ≈	TRI	rcr s	2 0	TRT Y	ATA	ATT
925 276	· 985 296	336		1165 356	376	1285 396	1345	1405	1465 456	1525	1585 496	1645 516	1705 536	1765 556

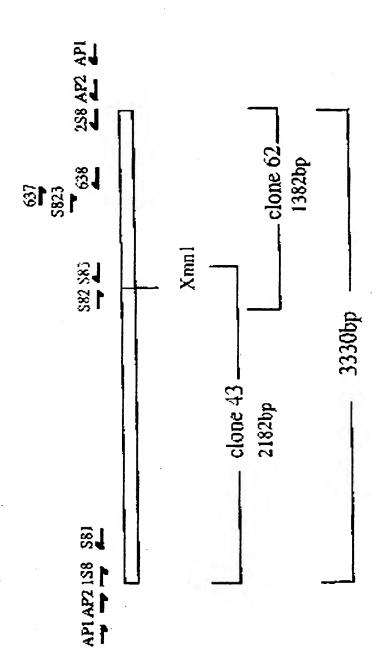
1884 595	1944	2004 635	2064	2124 675	2184 695	2244	2304 735	2364 755	2424	2484 795	2544 815	2604 835	2664 855	2724
S S	GCT	GCT	tt F	TIG L	A.F.G. K	(§ (%	AGT S	E CAG	rGF C	2 43 0	ATC I	ATA	r G	44 5 a
AIT	છું ∢.		7 กุล S	A F		CTA L	TIC	TTC L) (8)				CHA	CATA C
GTT	2 4 4	ភ ភូមិ	SCI A	3 ⊨	A C C	ರ ಗಾ	#G# \$	161 C	GAT	ATA I	85 O		CIT	T'T' (
7CT	**************************************	ည်	ata I	ATC I	KAT H	CTA A	A AG	CAC D	GTG V	ACT	D D	TIR	STG.	305
ATT	AAT	TTT	ည် ည		57.4 V	១៩១ ម	STC L	ភ្ជុ	TTT		TAT Y		Эд ±	S S
TTG	CIR	కే ప్రా	E W	TOG S	K K	GET o	TIC	ල අ	GRC E	or o	4 35 is	GTT.	CTG	æ
4 ×	\$ 500 s	្ត ១ ១	CAT O	10.00	S:3	2	rgt S	SCT A	CCC P	₹7.5 ¥	ტ ექე	450	A.3.G	CAT GGC AG H G •R ontinued)
A SA	TC1.	۶۵۲ د	្ស	కే గ	E E	 (5 to	CAT	cr: J	CCT *	8.4	કુકુકુકુકુકુકુકુકુકુકુકુકુકુકુકુકુકુકુ	ij k	223	CAT H Onti
TTS	CTC	TCC \$	X S	ii a	1. 1. 1.	្រូវ	ج ۳	6¢₹	GT¢	CAT	ង ប៉ូន្	ATC I	TTC	ည် ပိ
CTT L	CTT	RCT T	G Lagran	TTT	ATC H	5 ×	TCG S	eg CV	T&T ¥	A FOT	CAT B		TCB	ATC I
ACT T	A.S.	atc I	36G R	हुनु इ	ון היים	etr I	GCT A	G: 17	% 2€ 2€	SGT	113 1	J D	೨ ೨೨ ಕ	AGT S iguz
TCT S	. 663 6	CTJ	AIC	TTG	6) 2	ig F		- 35 a:	ž S	TCT S	ATT T	۶ چ	ర్దే 🚡	្ត ភូមិ
430 8)) (ATA A	GJC V	i S	े अर्थे अ	ង្គ	A AC	6C7	2 2 2 2	535 G	ACE T	AAG	GT& V	FC.
61G V	AAT) (3)	E .	4.00 4.00	37.0)) () ()	GTG V	K K	त् १३ स	ata 1	G&C D	G G G	i i i	6CT
ACT	ည်း (၁	ភ្ជា ភ្	GTT V	ATT	CAT D	3 ≠	ATT	11 13 17 13	TOT	P.A.C.	AAS N	Į Į	TCC	ני ט
ঠুৰ	GTT V	TTG	ec.	GCT ▶	61 7	Te.T Y	8	SAT	CTA 1	ATA I	83 a.	ATG	ig o	0 # 20
TCT S	GTG V	ည္ (၁)	<u>ل</u> 0	2 3) ()	8 8 8	i L	Ž.	ACT.	ට සිට ම	63C V	D Z	Grt	88
ដូច	Grt V	S P)) 1900	AAG *	CJG L	AIT I	CAT B	75.C Y	TCA S	GTT	TTC		SE C	7 T
e e	ال د د	S R	ដូ	A.A.G	e e e	ACT T	a G	1. AC X	CIT	ర్ట్ 🗸	A A	G G G B	GCT	TIC.
AGT	77.C.	CTT V	த ந்ப	ပ္က	ATA	A.A.G K	¥ H	A CH	350	GAR G	GAT	#CC .	ATG (GPT 7
1825 576	1985 596	1945	2005 636	2065 656	2125 676	2185 696	2245 716	2305 736	2365 756	2425 776	2485 796	2545 816	2605	2665 B56

2784 895	2844 918	2904 935	2964 955	3024 975	3084 995	3144 1015	3204 1035	3264 1055	3324 1075	3397 1082	3458 5	3522 16
TTA L	TJA L	ATR	ACA T	GTG V	٠ ا ا	AGA A	GAA E	G&C D	CJ3	TGA TITABICTIAACAITATAGCAACTGCAAGGICTTGAICAICTGTAGTIGCG	000 6	
ata I	ACF.	4 2 2	CTG	CTT	A CCT	ర్ట్ ఛ	អ្ន	GAG E	AGA R	ltrg	PAS N	5
GTT	ij Į	CCT	TAT	AAG	ATA I	CGT A	4) 4)	क्ष व	atc I	CTC	TCT 1	Mar
CIT	Trt Y	T Y	7 5 5	TRT Y	වූ ර	GT.A V	ည ဗ	TCT S	a a	ET CAST	GAT 7	A STAN
TCG	ರ್ ಬ್ರಕ್ತ	CAT K	GTC V	CTR L		GMG	AGA R	בים א	ეეტ ე	Tre	212	THE T
CGT &	A.T.F.	300	CS.3	7. 7.	() "J	್ರ ಕ್ರ	000 g	₩ \$	3CT A	GGTC	ANT THE AGAG A	IAA AALEEFAACEARARAARAARAAAAAAAAAAAAAAAAAAAAAAA
FCT	n U K	STC	TAC	ACC T	13 13 14 15	क् क्रु	ر در در	1.i.1	12.17. 1	7 809	رة م	44247744444444444444444444444444444444
TCT	इ.इ.	TTT	ACS T	GTG V	A CCI	TTG L	Çiri. E	2. A.	28 %	L) Ha	11 ·	ent:
161	ម្រិស	CTT	ც ც	GR.*	GHT:	K K	r F	CTG	TCG N	17360	13	क्षान्त्र श
) H	0.4c	GTT	GTT >	GRI	a Argon	35 ex	ర్ట్ _ల	ွှဲ မ	وري هوري	arte	∯ ப இ	c cca ima ad a Fignes 4
AGE A	g 3×	1133 L	TCT S	រួ ក	TTT	50 4	ig in	?; €	CSJ.	TRAC	AGA AEA R	(CC)
ATC	ACA	TGI	ត្តព្	GAT	504	₹ ₹ [5	ల్ల స్ట్రా	77.T F	AEG K	MTC.	ಸರ್ಣ ಸ ಸ	200 KOO
ata I	AGC S	A SH	S S S S	Tat ¥	FIT	ş v	ក្នុ ថ្ងៃ យ	TTG	TT.A. L	TIL	era ec 1	ಕ್ಟು
CAC E	PCT T	S X	7.7C	S & GI	3 S	at T L	র ও ও	G&C D	rif	ξ. 30	AGT HI S I	101 S
S S	ည် ပ ်	545 E		ည္သမွ	200	၁၁၁	atg K	GGT	GRG E	13.T		TTI F
909	AG!	SCA A	BAC	A A	S S	CGB R	3 2	É.	T'IC F	ភ្ជី	T ATT	ATC I
GAA C	66.9 6	CTB	AGT 3	TAS O	<u> </u>	S æ	ACA T	GCT A	₫ •	TGT	g TGT C	SEC.
AGT (CTT (21.3	ATC 1	X X	CTT	ATA I	FAR T	TCG S	Xes X	ACT	T ATG	2 2 2 2
TER I	GAG F	CAT (GAG 3	250	GGT	TGT	3.C.A.	ATT	TGG F	ర్ట్ .	tactaact	T IG
GRA 1	GAT C) <u>-</u> 0 0	17.00	TTG	55 ×	37.R	S S S S	TCT S	ن ان ان	¥ ¥	TACE	GTG V
2125 876	785 896	845 916	2905 936	956 956	1025 976	3085 396	1145	3205 1036	3265 1056	3325 1076	3398 1	3459 54

Figure 5

# 2 # 3 B C	~	S N N N N N N N N N N N N N N N N N N N
MESENTAL MESENTAL	(1) de (1) de	Burlaarerkorkorkovatovatovatovatovatovatovatovatovatovat
1000年		en benne ta appelletar abbut bet abbut total benne ta entropy of the collection of the state of the collection of
2000年20日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日本の日	200	it eare de socioque de Compartico de la compartico de la compartica del compartica de la compartica del
# # # # # # # # # # # # # # # # # # #	900	THE CORREST OF A SECRETARY SET OF THE VOICE SET SAGES OF SET OF S
25. 10. 10. 10. 10. 10. 10. 10. 10. 10. 10	77	efferences acellas en en en en en estados en Alboras estados estados en estados estados en
10年	44.18	OCHENTIONE SOND AND SET TO THE SET TO THE SET OF THE SE
20 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	5 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	er syksit vyvorker film kalika se olle i cles - sagatkia karakaska karakastio e in meska karaka ovsattulka Opiskim seksattur e piskim pavika kaka olle i cles - sagatkia karakaska kaka i karakaska karaka karakas ka
A NEW	# # # # # # # # # # # # # # # # # # #	vdskyjendykshikketskypeliktvaccaeskypelikatenskopiskespakiseskapenkateskapiskespiskeskissis Gredisesiskyskopiskopratikopaatonskyskiskoskopiskespiskiseskiskaioarkatesistaliskasissiskasis
96 T 4 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	7 0 0	Ceppella coofinded in thiony <u>vendta compert correcting contrades predessanded vendes a minau to competed bilaus</u> no abita roadinitudo - copusi <u>nishi</u> a un predessandes minacentes (resonales birghés par els labalturis en des competed son
27 C 25 Z	4. 4.	sea the season of the result of the season o
X & 8 8 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	7.9	en og stylken program i stylken styler so bronden beneden beneden beneden beneden beneden beste beste by stand De benes stylken beste beste som beneden stylken som beneden beneden beneden beneden beste stylken som beneden
MANALAC MANALAC MANALAC	1014	AM TA A M M A A A M M A A A A A A A A A

Figure 6



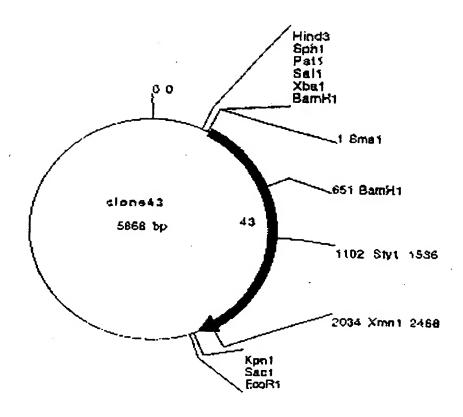


Figure 7

Communits/References: 43m 5' side of S8 (AIMSH6) 2182 bp in pUC18/Small

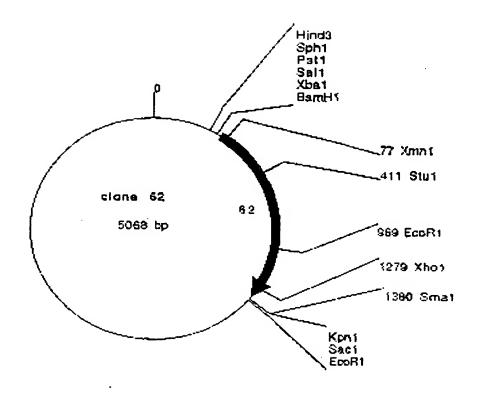


Figure 8

Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Smet

-	PARAG	AGTT	୍ଦ୍ରକୃତ	CCTG	ACCA	GTAT	CCTT	TCCG	CCAT	TTCL	ACCA	ದಿರಲಾ	1000	A. P.	ATTE	TCG	233	PATC	TTC	ttgagocotgaggatogtttoggocalttotaggaggraaagttlttggogggaaatttltgcococ	80
89 T	TTTCG	1 800	TTCT	CTC	CCIC	hatictcrccccaracalcottccccccccccccccacattccrara atg	CATO	GTTT!	CICT	S C C C C C C C C C C C C C C C C C C C	CTCT	CICIX	ACP.	ATTC(CAMA.	स ह इ	ଥିବ ଅ) # 9 9 0	ပ သ သ	၁ ၁ င	15.
154	4 20 4	3 S	ATT	25.7	FCT	TTC	TTC	CRA Q	A X	ည္သ	ACC 1	333 8	25.4	SI I	7 7	AAG K	GST	17G	GTT V	S	217
214	66C 6	GRT	SCT A	gCT ★	AGC S	ပ် ပြ	999	ပို့ လ	ပ္သင္သ ဇ	35.5C	§ 0	6 6	CC.	768 38	TTT F	2 K	GTG V	A.P.G.	55.55 54.55	86 G	27.54
274	CAT O	SCT T	Z ×	ပ္ပ ပ	ည်း ရ	GCT ₩	TCT	GTP.	CGT R	7 7.	100 a	GTT V	3 S	4 F	TCT \$	GTC V	GAT 0	GAG E	GTT V	A CA	3 4 4
334 65	5 0	ACG	GAT	ig .	5 ~	200	5 86	A.A.G	GTT v	۲ 200	CGT	CGT	ج ورا ت	ري د. ريو	ις 13 15		ğ _o	TIT	₹	00 04	9 8 4 8
394 85	GCT	GAG E	TCC	300 A	GST G	GAT D	ប្តីដូ	700 5	700 5	្ត ប្	TTC	35°	aat N	ait I	27G B	CRT REST	X K	TTT	GTA V		453
454	GTC	CAI	G&T O	S A	GAT	rer c			GAG E	35G R	ه. 5	§≈			5 5	S GTT C	ပ္တည	7	aat K	GAT	513 124
514	Į,	TCT	CTA L	101 C	ATG M	A A	000 A	4 1.5 1.			ATT I		နှင့် န		8 151	100 s	Z Z		S S	AAA K	573 144
574 145	F F	880 0	কু জু চ	AGA	R S	C&T H	GCT A	TTT E	e: S	17.0	r S	999	AGA X		5 5 6	CIT	AGA R	rca s	GTA V	GAJA E	633 164
634 165	GAT	ata I	გე ც	GTA V	GAT	ည လူ	GAT	GTT V	ដូច	0 0 0	800 P	4.00 13.00	ACA T	දී වූ බං	8 6	atg 8	R CG1	8 ~	CGT R	GCT	693 184
694 185	rcr s	C3C A	TTG L	A AG	5 %	GTT	CIG	34G 8	ÇAT O	44.9	atg B	FCT.	TTJ	X Xess	្តិ ម	GAT D	A AG	GTR V	CCT	6TB >	753 204
754 205	TTG L	a a a	S	A PAGE	A X	AGG R	01.3 19	3 ×	atg N	2 1 1 1 1 1	5 d	GR (900	GTT	16T C	56A G	GAG E	7. X	A.A.A.	GP.	813
814 225	GTP V	A.F.C	35 A	ថ្ងី ប៉ី ប	ACC	ዶችዶ ጽ	TTT E	ម មិន្តិ	TGG N	CTT	ର୍ଷ୍ଣ ଓ ଜ	s s	tcr	() () ()	U L	AGG.	GAT	8	rr X	AGA R	873 244
874 245	A L	CGT R	בכד ק	GAT D	GAT	884	CIT	TAC	o GAT	A GA R	s. K	ACC T	TTE [.	ည သိန္တာ	ATA I	CC P	CCT	GAT	GTT	TIC	933 264

Lynne

VSDOCID: <WO 9919492A2TI >

993 28¢	1053 304	1113	1173 344	1233	1293 384	1353 404	1413 424	1473 444	1533	1593 484	1,653	1713 524	1773 544	1833 564
SAC ATT D 1	TIR GGT L G	GIT GGT V G	AAA GTT K V	ACT AIN I I	ATC SGG I G	TGT TCA	TCC ATC	rag gar K e	TAT ACG	aca gat T d	GPA TCA E S	GCA GAG G E	ATT TTT I F	CIT GAG L E
RTG G	र्द्ध व	0 1 00	TAT Y	AAT N	P. P	RAG K	טפט טפט	CCA	A F	GAT	ICI	CTT	g o	AAT N
GRE TAT	GAT GOG D M	TGC REA C R	CST CGA	SCT GCT	043 643 643	CTA CM L O	TCG CTT M V	GIT ICT V S	CTA AGG	atc ccc n c	GGT TCT G S	Act cct s a	Cet GGG H	rtg gta N V
S. S.	CTA L	4.4.4 A. A	E 13 44	aca R	9 a	5 5 5 5 5 5 5	T T T E	S S S	: :: ::	GTA	A. N.	CTT L	₹ .	(9
CTT AAG	Trt Gre Y	676 658 V	CTA TTA L	434 GCC X	ACA GCA	म्हर्भ इ.ए. १८	TTG AGG L A		CA3 236 2 K	CCT CAS	tec tit Y	orr occ v	GIA CII V	AT GGC CAG AC G Q T (Continued)
kst s	בי בי	CCT C	25. 25.	GC:5.	년 () 년	ATC 1	50 مرد 10	ITA L	SCA A	GT3 V	18 18 19	Cat D	GAT D	ஷ் 🖙 🐸
TAT TGG X %	Tri Grg Y E	ATG AGT M S	GTG CAS y Q	ರಿಸಿದ ದಕ್ಕಿಸಿ ರಿ ರಿ	ACT CCA T P	Aga Gag X E	TGT GCT C A	00% GGG G A	aga gra k e	CCT CCA	TCT AAC S &	GAA TGT E C	cta gas L e	AGA ATT R I Pigure
50	rii F	D) #	6 5 4	S S	TTA L	atè I	GRC D	C:T	្តី ខ	rtc r	GA P	aat N	OH X	ם ב
א א א	GGG ARA G K	ars atg K b	GAT GAG 0 E	GRA ACA E I	CAG 6TA Q V	crr go	IIT GII F	GCT GCT	655 CTA 0	GTA CEG V Q	ata ata i i	GGT CITA G L	AGG CTA R L	cer rer G C
D s	ea GTG V	GAC TGG	36 ATA 1	CRG CTR Q L	CTA GTT L V	car crr n	rr gcr	TCA TGT	age ara	AUG GCG T A	aca aat r n	GIT CAT V D	CTG TCT	TRC RGG Y R
icr cap s A	TT REA F K	i L L	AGT GGG S G	ក្ស ភូមិ	AAG K	GTC	GGA TIT	5 a	ပ္စ (၁	TCT S	GTT V	GCT G	ក្ល ក	GII V
•	IT IIC F	4G GAG E	CT CAS	sa ayc I	Cib. AGG R	## GOC	rg TRT X	ат Сат О	TA TET	999 fo	CT 668 6	ac TGF C	I'X ANT M	AC CALA
AAC AAA ATG K K M	GTG C	CAC A	atc ti	0 45 0 8	ATT C	3 a	ACT G	사 다 교	GTG TTA V	TTG A	3) 44 1) 44	766 A	CTA A	CCA T
934 265	9\$6 285	1054	1114	345	1234 365	1294 385	1354	1414	1474 445	1534	1594 485	3654 505	1714 525	1774 545

1893 584	1953	2013 624	2073 644	41 13 8 66 4 3	2193	2253 704	2313 724	2373	2133	2493 784	2553 804	2613 824	2673	2733
A AC	GAT O	AGT S	S «	sts v	CAT E	AAA K	೨೨೨ ⊀	ACT	CAC	SCS.	Į,	GGT	att I	R
GAT D	¥ ¥	ट्रहें इं	နှင့် ၁	र्व इस्	att I		8 8	த் ந	att 1	GCT	A.P.A.	143	AGC S	<u> </u>
CII	CIC	ជ្ជ ភា	CTC	££3 K	\$:5 \$:5	CTC	र्ड ज	SCT *	\$ 67C	TCI	ර වූ ප	ე გ•	ဂ္ဂ (၁)	7
trt Y	A P	2 E	CTG J	000 0	kGT S	55 ×	77C F	# ₩	ម្កា ម	55.7	PAT N	<u>5</u> 4	AGT	ACT
₹ ¥	£2=	લ લુંગુ નઃ	Į,	ည် '3 ည	્યું જ જ	TAT Y	න් ප්	ទី ធ	TCT	AGT 5	CAG C	Cit V	3 \$	5 4
TAC	ក្តិ	ACG J	4 5 9	C:J	2'TC };	C T	TCT 8:	241 0	766 E	4) 4) 4)	SAT D	∯ 4	AGA R	X X
TTG L	A IC	۲ ابا ابا	778 1.	(; (; (;	ξ. Σ	7 (G	C:7 C	بر () غر	ජ වී ය	405 e.	¥08 1	Til	A CH	. 430 0.00 (best
ACC 1	76G	ಸ್ಟ್ರ ಕ್ಷನ್ನು	ည်း သ	٠ ن ئ	\$ % 5	λ:: S	T 17	STG V	t L	h 11	GCT.	<u>د</u> د	SCT *	000 000
556 6	rat. S	4. 4.9.03 3.03	ನ ಕ	:; ; ;	510 *	्र संभ	F. "	GÄT D	50 50 50 50 50 50 50 50 50 50 50 50 50 5	Į Į	443 143	CAT R	GAG E	ATG M
TC.	አርG	GTT V		SIG S	ATT:	ار ان ق	១មួច ឧ	S. O	છું જ જ	TTT	ş Ş	766 B	ပ္ပဲ	A.B.C.
5	T't L	स् ४.८७ ४.८०	ž ×	S	∯ (3) (3)	±3.5	STA L) 12 12 13 14 14 15 16 16 16 16 16 16 16 16 16 16 16 16 16	4 4 4 3 3 3	rct S	á Z	CTS L	CTL	CCA P Fig
ខ្លួ	17.0	5. 1	CAC 55	000 K	ري دي وي	10 8 8	500 5	ပ် ပို့		भेट्रेस श	S 4	4; 13 15 15	25.3	S S
ត្តិ ត	50 ×	C1.	ر 1:	ရှာ် နည်	TTT F	ور ن	الم الم	T.S.7 *	TTT	CTG L	ftt F	∯ 0 0	alta r	ACG T
647 0	A AG	ار ان چر	TRT Y	Į, n	ું જું ન	N N	A X	A Z	ניז	STC V	att K	ATC	GAT	CTG
ក្	567 6	A N	5 5 5 5	₹ 8 æ	49.X	3 a	ত জ	CC.A	633 6	GRI D	GTT	# F	A&T M	773 1
S S	ACT T	A.A.	0 0 0	GTT	stt V	CT.	GTA V	T'IT F	ATC.	CIA	CCT	CTT	9 2	TTG L
AAT N	g _	ATC	T T	သင္သ	g g	SCT A	ITA L	ე გე	CII L	ည် ကိ	AGG R	ATA	GIT	TCA S
AAC R	act s	S S	AIC	TCT	3 D	77G L	ata 1	3 8 8	ATA I	နှင်င) (2)	CCA P	223	99 æ
TTT	GTT	E E	మే చిత	X X	A A	TTG	CCT P	GAT 0	g K	r Tr I	atg K	9 <u>26</u> 6	7.T. J	P 2
ata I	1 1 1 1	GTR V	ATG K	ATC	CJC L	CTG	CIT	ata I	CTC L	F ACC	AGC		30	CAT R
3834 .565	1894 585	1954	2014 625	2074 645	2134	2194	2254	2314 725	2374	2434 765	2494	\$62 \$68	2614 825	2674 845

2793 884	2853 904	2913 924	2973	3033 964	3093 984	3153	3213 1024	3273 1044	3333 1064	3393 1084	3453 1104	3521 5	3579 19	3606
T TGC C	S ACA	A ACT T	a tac	C III	16 AAA K	o ChC	ar Gre V	CAA GCC	TCA AGT S S	AAT GCC N A	TCC TCT S S	CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT	AAA K	
1 1CT	ATG	S SCA	পুঠ্	5 CTC	7 7 CG	200 P	# O				¥	16 0	£ 3	
3 3	ATC	Z Z	GRT D	atg M	₽ 1	GAT	CTT	GCT A	TIC	ASC N	. AAA K	TTT	450 K	
TĜ₹ Ĉ	ACA R	30	TTC	ეეე გ	رئ •	TGT	წ ა	GCT A	945	T. C.A.C.	AIC	AGTC	រូវ	
ල් දුරු	O CAT	7	ACT T	JGI C	رن 13 ه	56T G	TAC	CGT	रुद्रम))) (*	943 9	GTT	TTT C	
GTG Y	rct s	5. 7.	AGT S	ණ. ට්රි ර	80	CG ₹	သို့ လ	ፓርሴ S	40a 8	GTC	C31	2225	TOC I	ন ব
7. 7.	ನ ಕಿ	င် လ ()	الم الم	str '	ე: ≃	ون دن م	ର୍ଷ୍ଟ୍ର ଅ	स ्छ स्टब्स	CJ'A L	# 55 %	TCG W	TG23	3. x	irme
757 0	ם ט	9334		Ź×	ากะ ห	3 0	CCT 5	#C# 1	သူ မ	TCT	TTR L	TATO	44.7 4.8	(Continued
ည ပိုင်	CIT	RCF T	40k R	ក្នុង ខ	8 8 8 8	TAT Y	1 6T ភ	ଖୁଣ୍ଡ ଅ	3. S	att I	76C	13 13 13	ATC A	. =
7 7	366 38	ស្លី ខ	ပ္သ ဗ	GTA *	77¢ ₹	GAT D	6€T \$	GT:	JCA S	୯ୈନ ଜ	TTT F	TAR	TAT A	igure 9
30	ACA Y	ACT	ยาว	CTG L	n A	र्टी इ	ភូវិភូ ភូ	() () ()	8 x	GTG V	TTG	#C#	T ATT L	u.
) } 4	TTC	76c	हा: क	CAC X	A DEA	A53	្ត ស្ត្	2.5.0 4.5.0	717	1. 1.	AC3	21'G	760 T	4 4 4
r t t	ATC I	GA. P. P. P	GAC D	೧GT R	# CC	ភ្ជី ។	A T) 14 z	AAC N	AD A	0 g	. €CT	T HIT X	3. R. A.
ATC 1	% 7 7	GTA V	CTJ	TTÜ	CTC L	₹ ×	TTA L	900 m	ಕ್ಷನಿ ವಿ	₽¥G K	TAC	, ATG	# 25. X	स्तुष्ट अ अ
GET V	CAT O	TTC	ATC I	GII	oct P	TTC	CG1	ata I	ე ე	្ត ស្ត្	0 0 0	T.	CTT	K K
$y_{\mathbf{a}}$	្វ ភូភិ	TTT	GTA.	77. S	S F	4 4 4	TAC Y	ი გეე). 1	7 2 2	ag o	\$ ×	CCT C	A.P.A.
CIG	ر. د تا	T	ig 13	TAC T	TAC	160 C	TTG	GCT A	្តិ ខ្	SAC O	द्भु ल	CCC	arr C	KAN X
TGI	သို့	AGT	ర్జ్ ఇ	ក្លី ខ្លួក	CAT H	Ç ₹	TTC M	ATG M	AGA A	હુ ગ ત્	် (၁)		ب	A.R.A.
5		980 980	GAT D	AIT	ACA T	atg M	इत्) 213 21	A.A.G	C to H	ATT	16T	RTG III M F	A A
	स्टु ज	র টু ড	95 o) (4	4 4 4	25 ≠	CTA	800 4	ATG 5	ព្	ည ည	TAC	æ E O	P.P.R.
734 865	794 885	\$5.85 90.5	2914 925	25 4.00 4.00 8.00	3034 965	3094 985	33.54	3214	3274 1045	3334	3394 1085	3454	3522	3580 20

Figure 10

- 1700 -	202		1 E P 0 44 - m1
1914 BYLESS CONTROL SON SON VERSON VIV	100		1 CH 3 44 - 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1016 British of the straightening of the control of	**		14 - 54 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Braner i Wie er in der ber in der ber ver von der bestelle bestellte bestell	7.0		10 4 4 4 5 4 6 5 4 6 5 4 6 5 4 6 5 4 6 5 4 6 5 4 6 6 6 6
A particular and the contraction of the contraction	X.F.		144
ALABERT CONTROL OF THE PROPERTY OF THE PROPERT		•	107 6 2 4 4 0 4
est interpretations of the contract of the con	22		78-92 9391
sviltossavata jupoumaes Balovaus (Rebust), vas (Astacoumes - otta is the same same validade tradition (Balo same seekvaluaes Estatora (Breille vasta vasta vasta respektivation (Breille vasta vasta) de vasta vasta	7 T		14-386 X31
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	777		14 - 54 4 + 5 1 4 6 1 5 1 4 4 5 1 5 1 5 1 5 1 5 1 5 1 5 1 5
oreo krittstra solitustavita estaktura kalendara proechi, proidera de sustanta sustanta ostata de vili de l'a Veresta de la russi presenta de la sustanta de la coubse de la sustanta de sustanta de la sustanta de la sustan		٠	100 4 A . 9 8 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
A PRIOTO GRAPH REPUBLIED NOT A SELECTAND FOR A PRIOTO PART CONTRACTOR SELECTED SELECTED FOR A SELECTION OF A SE	316	•	10 4 4 4 5 4 6 4 4 4 4 4 4 4 4 4 4 4 4 4 4
Dem ver ver ver gerinden de se samenten de ver de samenten de ver de samenten de samenten de samenten de des de de samenten de			78 - 48 7 9 8 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
マラザリエルのサイトには、「「「「「「」」「「」」「「」」「「」」「「」」「「」」「「」」「「」」「」」		7.4	

TOTALATARAAGRAMIGAAGGATATATTITTTTTCATTATCARG APARCARCAGEACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	TETTTGGTTGGTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA	50
ADAMCANCAGGETTTTTTTTTTTTTTACATTEGTCERCENANTACAN 150		100
CATANACGREATCHTTACCHTTTACCCAATTTTACCCCCTAATTACAA 200 CUTAGAACCTTCCCATCTCCCAATCTCCCAATTTACGAACACCTTT 250 MCCATTCCATATCCTCCAATCTCCTCAATCTCCTATTCTCTC 300 MCCATTCCCATCTCTCTCCCAATCCCCCTCATTCTCTC 350 ACCTCACCTCTCTTCTCCCAATCCCCCCCCCCACCTTTCTCT 400 MACCTCACCTCTCTTACACCACGATCCATCCCCCCCCACCCTTTCTCT 450 TICLAGCTCTTCAATCCCCGAACCCCTTTCCCCTTTTT 550 CTCACAGCTGCTAACCCCTTCCGCATCCTTTGCACCTTTTC 550 CTCACAGCTGCCAATCATATCACCCCATCCTTTTCCCCTTTTC 550 GGTATCGCTCGCAATCATATCACCACCACTCTTTCTCCCCTTTTCCCCTTTTCAACCCCCC		150
CUMARACETITECCAMETACCEGARGCAGCCTGATTAGGARCAGCTTT 250 ACCATTCECAMETACCTGATCCTCATTGATCATCT 350 CCAAATCCGGTTGTGACATCACTGATCTCCTCATTGATCATC 350 ACCATTCCCATTGTCATCATCATCTCTCTTCTCATCATC 350 ACCATTCCCGGTTGTGACATCTTCCCAATCCGGTTCTCTCTC		200
ACCATICTCATATICTGANCTACCTGATCCTCACATTETCTGTTTCG 300 ECAAATCGGCTTGACATCTTCTCCCAATCTCGGTTTCTGTATCATC 350 AACCTCACCTCTGCTTTCACAGGTCCATCGCCGCAGGGCTTGTTTCTTC 400 AACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGGCTTTTCTTC 450 TCCCAGAGCTGCATTACCCGTTTCACACCGCTACACTCTTTTCTCCTTTTTTTT	CATARACHACH (CATARCER) AGCACCCTRATTAGGAACAGCTTT	250
CCAAATCCCCTTGTACATCTTCTCCAATCTCCTTTCTATCATC 350 AACCTCACCTTGTTCACACGATCCATCGCCCGCAGGCCTTTTCTC 400 TICCAGCTCTGCTTACTCACCGAACCCCCCGAGCCCCTTT 450 TICCAGACCCGCATCCAATTCTTAACCGCTTTGAACCGCTTTTC 550 CTCACAGCCGCATCCAATTCTTAACCGCTTTGAACCGCTTTTC 550 CTCACAGCTGCGTTAACCGCTTCGGATCCCGTAGGTCTTTGTC 550 CTCACAGCTGCGTAACCGCTTCGGATCCCGTAGGTCTTTTC 550 CTCACAGCTGCGTAACCGCTTCGGATCCCGTAGGTCTTTTC 550 CTCACAGCTGCCTTAACCGCTTCGGATCAGTCTTTCTCTCTC	CCTAGAACUTTUTCUATUTTUBUARDUADAGOO CON TITO CONTROL CON	300
ARCCTCACCTCGGTTTCACACGATCCATCGCGCGGGGCTCTGTTTCTTC TTCCGGGTTTTTTACCGGTTAATCACCGGAACCGCGTTTTTTTT	ACCATTCTCATATTCTCAAACTACCTAATCTCCCAATCTCCCTTTCTGTATCATC	350
TICCAGCTICTIDETCTTARICACCOGRACCCCTAGATTTCCCTTT TGTTCGAACCGCGATCGCATTTCTTAACCGTTTTGAACCGCTTTTG TCTCAGACCGCGATCGCATTTCTGCATCCCGTAGGTCTTGCCTTTTTG TTTGATTTGTGAGAACTACTGGTTCCCAGTCTTGTTTTTTTT	CCAAATCCGCTTGTGACATCTTCTTCTCCCAATCTCCCTTCTTCTTCTTC	400
TGTTCGAACCGGCATCBAATTTCTTAACGGTTTGAACCGCTTTTTG TCTACAGGCTGCTTAACCGCTTTCGGATCGCGTAGGTCTTGGTCTTTTTG TCTACAGGCTGCGTTAACCGGTTTCGGATCGCGTAGGTCTTGGTCTTTTTTG GGTATCGCTCGGGAACTACCGGTTCGGATCGCGTAGGTCTTGTTACTCCCTG GGTATCGCTCGGAACTACCGGTTTGAAATTGAGAAACGAAAGGACAACGCGAAA ATTTTATTAATCTCGGTTTTGAAATTGAGAAACGAAATGGATTGACAAGA ATTGACACCATCTTTTTCTCCATTTTCGTTTCTGGAACTTCTTAGACATG ATTGACGACGATCTTTTTCTCCATTTTCGTTTTCGAACTTTTTTTT	AACCTCACCTCTGCTTYCACACGATCCATCGCCGCAGGGTCTGTTTT	450
CTCAGAGCTGCGTTACCGGTTCGGATCCCGTAGGTCTTGGTTTTG 550 TTTTGATTTGGAGAACTACTGGTTCCCAGTCTTGTTTACTGCTG 600 GTATCTGCTGGAGAACTACGTTGATTGAGAAAAGGAACAGCGCAAA 650 ATTTATTAATCTGAGTTTGAAATTGAGAAAGGAACAGCGCAAA 700 ATTGAGAGAGATTGAGATTGAGAAAGGATCAAGAATG 700 ATTGAGAGAGTTTTTCAGTTTTGACATTAATATAATAACAAGATG 850 ATTGACGACGTGCATTATCCAGTTTACCATTAACAATACCATTCTTTTTGGGT 850 ATTGACGACGTGCCATTATCCAGTTATCCAGTTAACATTAACATTAATA 900 AAAATTTCGTAAAATTATTTCGGTTCAATTGGGTCCATTACTTCAA 950 AAAATTTCGTAAAATTATTTCGGTTCAATTGGGTCCATTACTTCAA 1050 AAAATTTCGTAAAATTATTTCTGGTTCAATTCGATCCCTTTCCACAAAAAAAA	TICCAGCITCTICSTCTTARICACCOGRACECCCCIASACCCCCACACCCCTTT	500
TTTTGATTTGTGGAGAACTACTGGTTCCCAGTCTTGTGTTACTGCTCTG GGTATCTGCTGGGAGAACTACATGAGACAAAGGAACAACGCGAAA ATTTTATAATCTGATTTTGAATTGAGACAAAGGACAACAGACGAAA ATTTATAATCTGATTTTGTTATATATACGAAACGATGAAGATGAAGAATG TTGTTGAGAGGATTGTATATATA	TGTTCGAACCGGCATCGAATTCTTAACCGTTCAACCGCAACACCGCATCTTTTTG	550
GETATICTECTGEDATGTCGATGAATTERGRAGAAGAGAACAGEGGAAA 530 ATTTTATAATCCAGTTTCGAAATTGAGAAACGATGAAGATGA 700 TTGTTGAGAGGATTETGATATTATATACGAAGATTGATTCTTGAGAGA 150 ATTCGACATCTTTTCCCATTTCCCTCTCTGGAACGTTCTTTGGGT 800 ATTGAGACGTGTCATTATCCATTTGCACTTAACCATTGGTTCATTTGGGT 850 ATTGAGTCTTCGGTTCGATTTGGCTCAATTGGTTATATA 900 AAATTTCGGTACACATATTATCATTAAAATTAAGCTAACCA 950 AAATTTCGGTACACATTATCCGATTCCAATTCAGATCAATAACCA 950 AAAATTTCGGTACACACGATATCCAATTCAGATCAATACCACA 4000 AAAATTTCGTAAAACTCGAGTCCAACCGAAAACCAATCCACACACCACAAAAACACTTCAACAATTCAATAAAAAA	CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGGTAGGTCTTGGCTTTTTTTCTCCTC	600
ATTITATIANTECAGTITICAAATTGAGAAACGATGAGATGAGATG TTGTTGAGAGGATTGCATATTTATATACGAGAATTGCTTTCTGGAGA ATTCGATCATCTTTTCTCCATTTCCTCTTCTCTCTGAGATTGCTTTCTGGAGA ATTCGATCATCTTTTCTCCATTTCCTCTTTTGCACTTCTTGGAGATTG ATTGGATGACGACGTTCATTTTCTATTTTGGTTTATACTATTTTTTGGTT TGGATCGTTGGTAGACATATTTCCGATTTTGGTTAAATGCTTTTTTTGGT ATTTGGTTTTCGGTTCGG	TTTTGATTTGTGGAGAACTACTGGTTCCC.AC.TC.TGTSTTACTSCSCSCS	650
TIGHTCAGAGGATTETCATATTATATATACGAAGATTGGTTTCTGGAGA 300 ATTCGATCATCTTTTCTCCATTTTCCTCTTTGGAACGTTCTTTAGACATG 300 ATTGACGACGTGTCATTATCTCATTTTGCAGTTAAACACTTTTTTGGGT 350 ATTGACGACGTGTCATTATCTCATTTTGGAGTTAAACACTTATATATA	GETATCTGCTCGECATCGTCGATGAATTGAGAGAAAGGAACAACGCCAACA	700
ATTECATCATTTTCTCCATTTCCTCTCTCTCGAACGTCTTTAGACATG ATTGACGACGTGTCATATCTGATTTGCAGTTAGCAGTCTTTTTGGGT TGBATTCGTGTACACCATATTATCCGATTTGGCTCAATGGTTTATAATA AATTTGGTTTTCGGTTCGG	ATTTTATTANTCTCAGTTTTGAAATTGAGAAACGATGAAGATGAAGATG	750
######################################	TTGTTGAGAGGATTETGATATTTATATATACGAAGATTGGTTYC TGGAGA	aco
######################################	ATTCCATCATCTTTTTCCCATTTTCCTCTCTGUAACGTTCTTAGAGATG	850
TOSATTCGTGGTACACCATATIATCGCATTTTACCTCAATGCTTTATATA AATTTGGTTTTCGGTTCGG	ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTTGGGT	
AATTTGGTTTCGGTTAGATTATGATTAAATTAAATTAAA	TOGATTCOTGGTACACCATATCACCATTTGGCCTCAATGGTTTTATALA	
AAAATTTTCGTAAAATTTATTCGGTTCAATTCGGTCCATG 1050 GAACCGAATTATTCGAAACGAGGGTTATCCGAACCGAATACCAATGCCTG 1100 ATTGACCGTTGGCTAGAAAGATCCAACGGTTAACAATAATAAAAAGTT 1100 ATGGGACGGTATCAAAGAGTCAAAGAGTAAACAATAATAAAAAGTT 1200 GAGCCCTUAGAGTTCCAAAAAGTTACAAAAAAGTTACAAAAAAATTCTAAAAAACATC 1250 TTTTTGGCGCCAATTCCCACAATTCCAAAAAATGCAAGCGCAAAAT 1300 GTTTCTCTCCACCTCTCCCACAATTCCAAAAAATGCAAGCGCAGAGAT 1350 CGATTTTGTCTTCTCCAAAAACCCAACGCGGGACCACGAGAGGTTTG 1350 CGATTTTCTCTCTCCAAAAACCCAACGCGGGGACCACGAGAACCACGATT 1400 AATGGAAACGAAACGCAACGCAGCAGAACCACGATTT 1490 AATGGAAACGAAACGCAACGGATACCACCAGGAAACCACAGGT 1500 TTCGAAATCTTCCAACAAGGAAACCACACGGATACCACCAGGT 1500 TTCGAAATCTTCCAACAAGAAAAAAATCACCACCAGGATACCACCAGGT 1550 TTCGAAATCTTCCTCCAGCAATATTATAGCAATAAAGTTCCACCAGGT 1600 TTCGAACGAGAATGTTCTCTGAATAAACTTCCAAAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAAATTTCAAAAATTTCAAAAAA	AATTTGGTTTTCGGTTCGGTTATGAGTTATCATTATAAATTAAGGTAACCA	7
GAACCGAATTATTCGAAACCGGGGTTATCCCCAACCGAACCAACC	AAAATTITCGTAAAATTINTTTCGGTTTCAATTOGGATCCCTTACTTCCA	•
ATEGGACGGTCATCAAAGCCTCAAAGAGTBARCAGTCAACAAAAAGGTT GAGCCCTCAGGAGTATCGTTTCCCCCCATTTCTACCACCAAGAGCGAAAAT 1250 TTTTGGCGCCAATCTTTCCCCCCCTTTCGAATTCTCAGCTCAAACATC GTTCTCTCTCTCTCTCTCACAATTCCAAAAAATGCAACGCGAAAAT 1350 CGATTTTCTCTTCTCCAAAAACCCACGGGGGGACTACGAAGGGTTTG GTTCCGGCGATCCTCCTACAAAACCCACGGGGGGACACCACCATTT 1490 AATGGAAGGAAGGGGATGCTAAAGGCGACGGCAGGGAACCACCATTT TCGAAATCTGTCCATGAGGGTAAGGGGAACGGATACCCCCGGGAAACCACCATTT TCCGAAATCTGTCCATGAGGTAAGGGAACGGATACTCCCCCGGAAAACCCACGGT TCCGCGTCCTGTCCT	GAACCGAATTATTCGAAACCGGGGTTAGCCGAACCGAATACCAATGCCIG	
TTTTGGGGCCAATCTTCCCCCATTCTACEACGCAAGCGAAACATC TTTTGGGGCCAATCTTCCCCCCATTCCAAATCTCTCACCTCAAACATC GTTTCTCTCTCTCTCCACAATTCCAAAAAATGCAGCCCCAGAGAT 1350 CGATTTTGTCTTTCTTCCAAAAACCCACGCGCGCGCACTACGAAGGGTTTG GTTTCCGGCGATCCTCCTAGCGCGCGGCGCG	ATTEACTCGTTGGCTAGAAAGAYCCAACGTATACAATAATAGAACAA	1150
TTTTGGGGCCAATCTTCCCCCTTTGGAATTCTCTCAGCTCAAACATC GTTTCTCTCACACTCCCCACATTCCAAAAAATGCAGCGCCAGAGAT 1350 CGATTTTGTTTTCTTCCAAAAACCCACGGCGGCGCACTACGAAGGGTTTG GTTTCCGGCGATCCTCCAAAAACCCACGGCGGCAGCGGAGACCACGATTT AATGGGAAGGAAGGGATGCTAAAGGCGACGCTTCTGTACGTTTTCCTGT TTCGAAATCTGTCGATGAGGTAAGGCGACGCTTCTGTACGTTTTCCTGT TTCGAAATCTGTCGATGAGGTAAGAGGAACGGATACTCCACCGGGAAAGG TTCCGCGTCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT GATCGTTCCTCCCTGTCTTCCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGAATTGTTCTGGAGAGAGAGATCTTCTTCGATTCTCTTAATTT 1650 TGTTATCTTTAGCTGGAAGAGAAGAATTCTTCGATTCTCTTAATTT GEAGAGATTCTGATTACTGCATTGGATCGTTTTTACAAATTTCAGCAG CCGAGAAGATTCTGATTACTGCATTGGATCGTTTTTACAAATTTCAGCAG CCGAGAGATTCTGATTCCGCTGAATGATTCATCTCTTATGATGAAGGCTA AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTAAAAGGCTACAAAAAGA 1950 TATAGCAGTAGATGGCGATGTTCTTGGAAACCAAGGGAACCAAATTG CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAAGAACACAAGGAACCCAAATTG GAGGATAAGGTTCCTGTATTGGACTCAAAAAGGCTGAAAATGCTCCA CACGTGCTTCTCGCTTGTATGGACCTAACAAAAGGCTGAAAATGCTCCA CACGTGCTTCTCGCTTATTGGACCAAAAAGGCTGAAAATGCCTCCA CACGTGCTTCTCGCTTATTGGACCAAAAAGGCTGAAAATGCCTCCA CACGTGCTTCTCGCTTATTGGACCAAAAAGGCTGAAAATGCCTCCA CACGTGCTTCTCGCTTATTGGACAAAAAGGCTGAAAAATGCCTCCA CAATCCGGTTTCTGGAGAGAAAAAAAAGGCTGAAAAATTGC CAATCCGGTTTCTGGAGAGAAAAAAAAGAAGTAAAAAGGCTCCTGAT AATGGCTTCTCCGAATCAGGGAAACACAAAAGGCTCCTGAT AATGGCTTCTCCGAATCAGGGAAACACAAAAGGCTCCTGAT	ATCGCACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAA	1200
GTTTCTCTCTCTCTCTCTCAATTCCAAAAATGCAGCCCAGAGAT CGATTTTGTTTTCTTCCAAAAACCCACGCGGGGACTACGAAGGGTTTG GTTTCCGGCGATCCTCCTAGCGGCGGCGACCACGCGGAGACCACCATTT AATGGAAGCAAGGGATGUTAAAGGCGACGCACCACGAGAGACCACCATTT TTCGAAATCTGTCGATGAGGTAAGGCGAACGGATACTCCACCGGAGAAGG TTCCGAAATCTGTCCATGAGGTATAGAGGAACCGATACTCCACCGGGT GATCCTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGAATTGTTCTGGAGAGAGGATACTCTTCTAATTT GGACGAGAATTGTTCTGGAGAGAGGATACTTCTTCTTTAATTT TTTTTTTTTT	GAGCCCTGAGGAGTATCGTTTCCCCCCATTTCTACGACGCCAAGGCCAAAAA	1250
CGATTTTGTCTTCTCCAAAAACCCACEGGGGGACTACGAAGGGTTTG GTTTCCGGCGATGCTECTAGEGGCEGGCGCAGCGGAGGGATACCACEATTT AATGTGAAGGAAEGGGATGCTAAAGGCGACGCATCCTTTTTCCTGT TTCGAAATCTGTCGATGAGGAACGGATACTCCACCGGAGAAGG TTCCGAATCTGTCGATGAGGAACGGATACTCCACCGGAGAAGG TTCCGCGTCGTCTCTCCCAATATTATGCATAAGTTTGTAAAAGTCGA GATCCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAAGAAGAAGATTCGTGTAATTTTCAAAATTTTCAGGAG GAGAGATTCTTTAGCTGGAAGAAGAAGATCGTTGTTTACAAATTTTCAGGAG CCGAGAAAGATGTTCTCCCTGAATGATTCATCTCTATGTATG	TTTTGGCGCCAATCTTTCCCCCCCTTTCGAATTCTCTCAGCTCJOCACATC	1300
GTTTCCGCGATCCTCCTACCGCCGCCCGCAGCAGCGCAGCACCACTTT ANTGTGAAGCAACGGATGCTAAAGGCCACCCTTCTGTACGTTTTCCTGT TTCGAATCTGTCCATCAGAGGAACGGATACTCCACCGGAGAAGG TTCCGCGTCCGTCTGCACTTAAAGCCGACTCCACCGGAGAAGC GATCCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGGTACTCATCTTCGATTCTCTAATTT TGTTATCTTTAGCTGGAAGAAGAAGATCCGTGTAATTTTCATTTCAGTAGTTTCTTTAGCTTTAGCTTGAATTTCTTCGATTCTTTAAATTTTCAGCAG CCGAGAAGATGTTCTGCATTGGATCGTTCTTTACAAATTTTCAGCAG ACCAGAGATTGTTCCGCTGAATGATTCATCTCTATGTATG	GTTTCTCTCACTCTCTCTCACAATTCCAAAAATTCCAGUUUUAAAAAT	1350
ANTGIGNAGGAREGGATGUTANAGGGAREGUTTUTGTAGGTTTTCCTGT TTCGNAATCTGTEGA7GAGGTTAGAGGAACGGATACTCCREEGAGAAGG TTCGNAATCTGTEGA7GAGGTTAGAGGAACGGATACTCCREEGAGAAGG TTCCGCGTCGTGCCTCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT GATCGTGTCCCCTGTTCTCCAATATTTTCCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGGAGATACTATCTTCGATTCTCTTAATTT 1000 TGTTATCTTTAGCTGGAAGAAGAAGATCGTTGTTTACAAATTTTCAGAAG CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	CGATTTTGTCTTTCTTCCAAAAACCCACGGCGACGACGAAGGGCTTCC	1400
ANTGTONAGGARGGGATGUTARAGGGGAUGCTTCTGTAGGTTTTCUTGT TTCGNAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG TTCCGCGTCGTCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTCTTAATTT 1700 TGTTATCTTTAGCTGGAAGAAGAAGATCGTTGTTTACAAATTTTCAGGAG CCGAGAAGATGTTCTCCGCTGAATGATTCATCTTCTATGTATCAGAGCTA ATGATGTTATCCTCAATTTCGTTCCAATAATGGTAAAACTCAAGAAAGA	GTTTCCGGCGATCCTSGTAGCGGCGGCAGCAGAGAGACACCATTT	1450
TTCGNARTCTGTEGATGAGGTTAGAGGARCGGATACTCCACCGGATACTCCACCGGT TTCCGCGTCGTGTCTCCCGCTCTGGATTTAAGCCGGCTGAATCCGCCGGT CATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGAGTACTAATCTTCGATTCTCTTAATTT TGTTATCTTTAGCTGGAAGAAGAAGATCCGTGTAATTTTCATTTCATTT GEAGAGATTCTGATTACTGCATTGGATCGTTTTTACAAATTTTCAGGAG CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	AATGTGAAGGAAEGGGATGCTAAAGGCEACGCTTCTGTACGTTTTCCTGT	1500
GATGETTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGGTACTARTCTTCGATTCTCTAATTT TGTTATCTTTAGCTGGAAGAAGAAGAAGATCGTTGTATTTGTTTTTTTT	TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG	/
CATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA TGATCGAGATTGTTCTGGAGAGAGGTACTARTCTTCGATTCTCTAATTT TGTTATCTTTAGCTGGAAGAAGAAGAAGATCGTGTAATTTGTTGTATTCGTT GGAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAAATTTTCAGGAG CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	TTCCGCGTCGTGCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	1600
TGATCGAGATTGTTCTGGAGAGAGGTACTARTCTTCGATTCTCTTAATTT TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTGTTTTCGTT GEAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTCAGGAG CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	- · ·
TGTTATCTTAGCTGGAAGAAGAAGATCETGTAATTTGTTGTTTTCGTT GEAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAEGAG CCGAGAAGATGTTGTTCCGCTGAATGATCATCTCTATGTATG	7GATCGAGATTGTTCTGGAGAGAGGGTACTAATCTTCGATTCTCTTAATTT	•
GEAGASATTCTGATTACTGCATTGGATCGTTGTTTACARATTTTCAGEAG CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATG	TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTGTATTCGTT	
CCGAGAAGATGTTCTCCCCTGAATGATCATCTCTATGTATG	Geagagattctgattactgcattggatcgttgtttacaaattttcaggag	
ATGATETTATTCCTCAATTTCGTTCCAATAATGETAAARCTCAAGAAAGA AACCATGCTTTTAGTTTCAETGGGAGAGCTGAACTTAGATCAGTAGAAGA 1950 TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACADCAGGGATGCGTC CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG CACGTGCTTCTCGCTTGAAGCGAGTTCTGAACAAAGGCTGAAAATGCTCCA CAGGCGTTTGTGGAGAGAAGAAAGAAGTAAAAGGCTGAAAATTTG CATGGCTTTGTGGAGAGAAAAAAAAGGAACCAAATTTG 2150 AATGGCTTGAGGTCTTCTCGAATCAGGGATGCCAATAGAAGAGTCCTGAT	CCGAGAAGAYGTTGTTCCGCTGAATGAYTCATCTGTATGTATGAAGGCTA	
AACCATGCTTTAGTTTCAGTGGGAGAGCTEAACTTAGATCAGTAGAAGA TATAGCAGTAGATGGCGATGTTCETGGTCCAGAAACACCAGGGATGCGTC CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGGATGAAATGACTTTTAAG CACGTGCTTCTCGCTTGTAATGGACCTCTAACAAAAGGCTTGAAAATGCTCCA CAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTTGAAAATTTG CGATCCGGTTTGTGGAGAGAAGAAAGAAGTAAACGAAGGAACCAAATTTG AATGGCTTGAGGTCTTCTCGAATCAGGGATGCCAATAGAAGAGCTCCTGAT	a 163 tottattcctcaattcctclaatactaatagetaaarctcaagaaaga	<u>- </u>
TATABGAGTAGATGGCGATGTTCTTGGTCCAGAAACADCAGGBATGCGTC 2000 CACGTGCTTCTCGCTTGAAGCGAGTTCTGBAGGATGAAATGACTTTTAAG 2050 GREGATAAGGTTCCTGTATTGBACTCTAACAAAAGCTTGAAAATGCTCCA 2100 GGATCCGGTTTGTGGAGAAGAAAGAAGTARACGAAGGAACCAAATTTG 2150 AATGGCTTGAGGTCTTCTCBAATCAGGGATGCCAATABAAGACGTCCTGAT	ARCCATGCTTTTAGTTTCRGTGGGAGAGCTGAACTTRGATCAGTAGAAGA	•
CACGTGCTTCTCGCTTGRAGCGAGTTCTGGAGGATGARATGACTTTTARG 1050 GREGATRAGGTTCCTGTATTGGRCTCTRACARARGCTTGRARATGCTCCR 1000 GGATCCGGTTTGTGGRGAGRAGRARGRAGTRRACGRAGGARCCARATTTG 1500 ALTGGCTTGAGTCTTCTCGRATCRGGGRTGCCRATRGRAGGCTCCTGRT	TITAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC	
GREEATAAGGTTCCTGTATTGEACTCTAACAAAAGGCTGRAAATGCTCCA 2100 GGATCCGGTTTGTGGAGAAGAAAGAAGTRAACGAAGGAACCAAATTTG 2150 BATGGCTTGAGTCTTCTCGAATCAGGGATGCCCAATAGAAGACGTCCTGAT 2200	. Cacgigcy totogettgragegagiyetggaggatgaaatgacittyaag	
GGATCCGGTTTGTGGAGAGAAGAAGAAGTRAACGAAGGAACCAAATTTG 2150 AATGGCTTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	· GASSATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA	
ANTEGETTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	CCATCCGGTTTGTGGAGAAGAAGAAGAAGTAAACGAAGGAACCAAATTTG	
GATCCOCTTTACGATAGAAAGACCTTACACATACCACCTEATGTTTTCAA	ALTEGETTGAGTCTTCTCGAATCAGGGATGCCPATAGAAGACGTCCTGAT	
	GATCCCCTTTACGATAGARAGACCTTACACATACCACCTEATGTTTTCAA	2100

GARARTGTCTGCATCACARAGCARTATTGGACTGTTARGAGTGARTATA	2250
TGGACATTGTGCTTTTCTT7AAAGTGGTTAGTAACTATTAATCTAGTGTT	2300
CARTOCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2350
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAATTAG	2400
GTCACAAGGAGCTTGACTCGAAGATGACCATGAGTCGTGTGGGAAAATGC	2450
AGACAGGTAAATTAGSTGAAACAACTGGCCTGCTTGAATTASTGTGTCTA	2500
TAARTTTTGACACCACCTTTTGTTTCAGGTTCGTATCTCTGAAAGTGGGA	2550
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	2600
actitatggaattegtitactgetacttcggctaggatttaagaaatgga	2650
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTE	2700
TTATTEGTTATATAATAAGTETTTCATCAAGGAGTAACAAACATAACTT	2750
BCACAATATTGTGTGTCACTGGCAAGGCATATATACCCAGCTAACCTT	2800
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	2950
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTGAAGCCTTGTGTTGTAAA	2900
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATG	2950
CTTTCETATTCTACTAATTATGTATTGACACACACTCGTTGTTTCTTTC	3000
AGATATAAAGTTGGACCAATCCAGCAGCTAGAAACATCTGACCAACCA	3050
AGCCAGAGGGCCANTACTCTAAGTTTCCTTGGATAGGCCAAAGAGAGAGAGTC	3100
TTGCAGACTGTTTTTCATCATTTTCTTTTTTTTGTACATTACTTTCATGCTG	3150
TARTTAACTCARTUCCTATTCTGGTGTGATTATCAGATAATTCCAAGGAA	3200
GCTAGTTCRGGTATTRACTCCATCARCAGCAAGCGAGGGAAACATCGCGC	3250
CTGATGCCGTCCATUTTCTTGUTATAAAAGAGGTTTGTTATTTACTTATT	3300
TATOTTATCATGTTCAGTCATCCAAGTCCTGAAAAATTACACTCTTCTT	3350
TACCARTCTTCCATCAAGCTGTGTAAAGGATTTGGAATAAGAAAATCATT	3400
ATTTGATGCTTTGTTTTATATGCAAGAGGTTCCCTTGAAAAGATCTGTTT	3450
AAGATTETTTGCACTTGAAAAATTUAATCTTTTTAAGTGAATCCCCTACT	3500
TTCTTRCAATGRTCATAGTCTGCAATTGCATGTCAAGTRATATCATTCCT	3550
*STTACTSCATGCCCCTCTTICTTAATGACCATTGTCTATGTTGTCTTTG	3600
TCTCGTGTGCTGGAGAAAATGATAGCTGATCCAAGCTCTACATTA1'CATG	3650
ATTAXGTAGCTGETCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG	3700
ATGTCAATTTTCTTCTGRATCTTTATTTTAGATCAAARTGGAGCTRCAA	3750
ANGIGITCANCISTSTATGGATTTGCTTTTGTTEACTGTCCTGCCTTGAG	900E
GTTTTGGGTTGGCTCCATCAGCGATGATGCATCATGTCCTGCTCTTGGAG	3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTATGTGTACCATG	3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCACCA	3950
TTTGGTATGTTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA	1000
AAGGTAAACTECTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT	4050
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC	4160
TATCANGAGAACCACAAAGGCTETAAGGAAATATACGTTGACAGGTACC	4150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATAGG	4200
TETETTARACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG	1250
GTCTACEGCGGTACAGTTGGCTCCAGTACCACAAGTAATGG66GATACAG	4 300
atgetgetggagttagaaatataatagaatetaacggatactttaaaggt	4350
TCTTCTGAATCATGG&ACTGTGCTGTTGATGGTCTAAATGAATGTGATGT	1400
	•

. I

	.150
1gcccttagtgc1ct1ggagabetaattaatcatctgtc1aggctaaagg	4450
TOTOTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4500
TTC&T&ACTTACGGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4550
ayggggatatítttccataccaagittacaggggttgfc7cagaaitgat	4600
GCCCAGACGATGGTAARTCTTGAGATATTTRACAATAGCTGTGATGGTGG	4650
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA	4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
agccgatttttgtttacctatgtagggaccttgtacaaatatcttgataa	4900
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC	4850
Cactcaragatgtagaaagcatcaataaacggcttgatgtagttgaagaa	4900
Tycacggcapactcagraagtatgcaaatcactggccagtatctccacaa	4950
acticcagactiagaaagactectccgacgcatcaagictagccttcgat	5000
CATCAGCCTCTGTGTCTGCTCTTCTGGGGAAAAAGTGCTGAAACAA	. 5050
CGAGTAAGTATCAATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT	5100
ATAGGACTARARCATTRCGGGTCTACCTARAGRCTCTTCTCCTTCTTTTG	5150
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5200
AARGERTTTGGGCAARTTGTGRAAGGGTTCRGAAGTGGAATTGRTCIGTT	5250
STTGGCTCTACACAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT	5300
TOTTTTTTTTTAGGANTXIUUUDAAAAAGUAATTATTTTTTTTTTTTTTTTTTTTTTTT	₋ \$350
CAATTCGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTCCCCAT	ू 5400
CTATCTTTCATACTTTACAACAAAATGTCTCTCACTACTCAAAAGCAATGC	<u>t</u> 5450
ATATGGCTTAGATCTCAACTCACACCCCGAGGATCCTAAAGGGATTTGCT	5500
TTTTATTCCTRATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT	5550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAAACCCCTGAAACTC	5600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGA&	5650
GTCATTERCACCATARCCTGCCTAGATGTCCTGAGATETTTTGCARTCCC	5700
AGCARGTCTCTCTGCTGCAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5750
CAGAAGCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC	5800
CAAGGACTATGGCATCCATTTGCAGTTGCAGCCGATGGTCAATTGCCTGT	5850
7CCGRATGATATACTCCTT6GCGAGGCTAGAAGAAGCAGTGGCAGCATTC	5900
ATECTEGGTCATTGTTACTGACEGGACCAAACATGGGCGGAAAATCAACT	5950
CTTCTTCGTBCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6000
dgitagataattactetattetttetetgeaatcagttctteaacatgaataat	6050
PAATTCTGTTTTCTCTGCAGCTTGGCTACGTGCCGTGTGAGTCTT	6100
GCGARATCTCCCTCSTGGATACTATCTTCACAAGGCTT6GCGCA5CTGAT	6150
AGAATCATGACAGGAGAGAGTAAGTTTTGTTCTCAAAATACCAATTCCTC	6200
GAACTATT7ACTCAGAT177GTCTGATTGGACAAGGTGGTTTTGCT77TT	6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT	6350
AGTACTTTCGATGGATACGECATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCARCTTATACTTGTTGATCARCARARACATGCARTTCATTTTGCTGAR	6450
ACTTATTGATTTATCAGGTTTTTCGTCACCTGGTAGAGAAAGTTCAAT	6500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCROCCACETGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG	6600
Figure 11 (Continued)	35 15 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

atctgattatcaaccacgtggttgtgatcaagacctagtgttcttgtacc	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	670D
atggetggaataccaaaccaagtggttgaaacagcatcaggtgctgctca	6750
AGCCATGRAGAGATCAATTGGGGAAAACTTCAAGTCRAGTGAGCTAAGAT	6800
CTGAGTTCTCAAGTCTGCATGAAGAC7GGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTEGECCACAACAATGEECCCCATTGGCGAAGATGACTACGACAC	6900
TTTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	695D
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAATTGTTTCC	7050
TCGATTATAACAAGATTATATGTATCTGTCGGTTZAGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA	715D
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATRACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT	7250
Tataagacttaagttctgaattccctaaaaccaaaaaaaaa	7290
TATTTTGTTGCATCTACAARCAAACCTACATAGTTTATAACTTAC	7350
TCATCACTGAGATTAACATCAGAATCATTCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTTCACGTAACC	7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCRACTCC	7500
AAATCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	7600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG	7650
TACTCAACCTCATCATCCTCATCCTCCTCCTCTCAGAATCACCAGGACT	7700
ATCCATCATCATCAAACCCRTTAGACTTATCTAAATAAACCTTAGTGT	7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	7800
GACTTEGGOGARACACARAGCARCATATCCRACTTATTGARARACGACCA	7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTTCTAT	7900
ACTICCTCTTCAAGTCATCATCATCTCCTACATTGCGTCTCAGATTTC	7950
TECATCETTAGETECTCACTCACTTTETCAGCTACTTCATTCCAATCETC	8000
GTTCCTCAAACTCCTTCTACCCAATTGCRAAAACCTATCTCCCCAAACTT	8050
CARGCAACACAA	8050 8062
	909Z

Figure 11 (Continued)

DOCID: <WO__ 9919492A2TI >

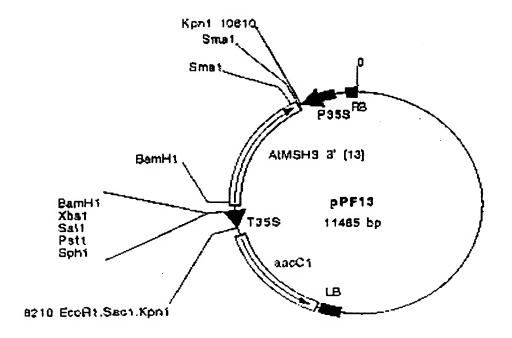


Figure 12

Comments/References: AtMSH3 3' side antisense : AtMSH3 3' (19 = 2104bp) from pUC18/13 Sal1/Set1/T4 Into pCW184 BernH1/T4 In Agrobacterium LBA44O4

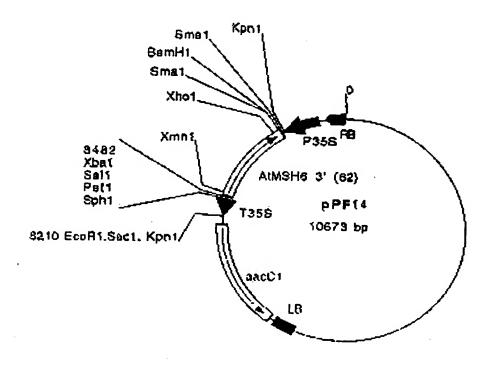


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sat1/T4 (1379bp) into pCW164 BernH1/T4

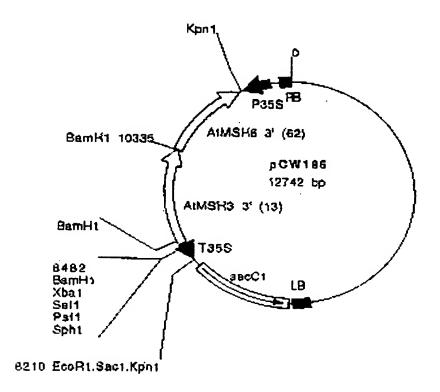


Figure 14

Comments/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (S8) 3' side (62=1379bp) Salt/Satt/T4 Into pPF13 (pCW164 AtMSH3 (S5) 3' side (13=2104) antisens)/Smat. in LBA4404

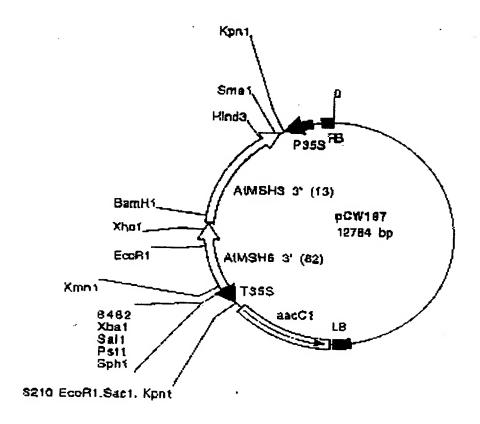


Figure 15

Comments/References: AtMSH3 3'/AtM8H6 3' antisans (0): AtMSH3 (35) 3' side (19=2104bp) Sal1/SstVT4 into pPF14 (AtMSH8 (35) 3'side (62=1379bp) antisanse into pCW184)/Small in LBA4404

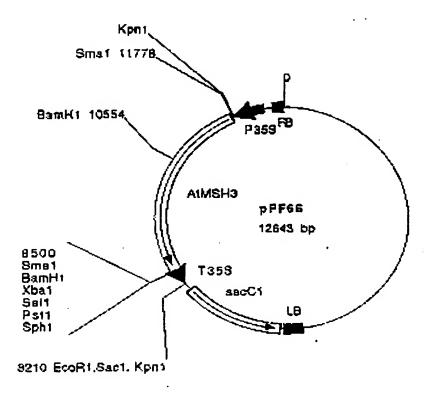


Figure 16

Commente/References: AtMSH3 (SB) complete, sease orientation : pPF28 (3342bp) Small into pCW164 Small

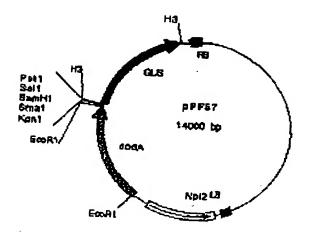


Figure 17

Comments/Reforences: pPZP111 with code ExpR1 cases in Score sile and Hinds GUS cases in Hinds elle. Kenfl. All gapes under Promoter/Reminister ass

Figure 18

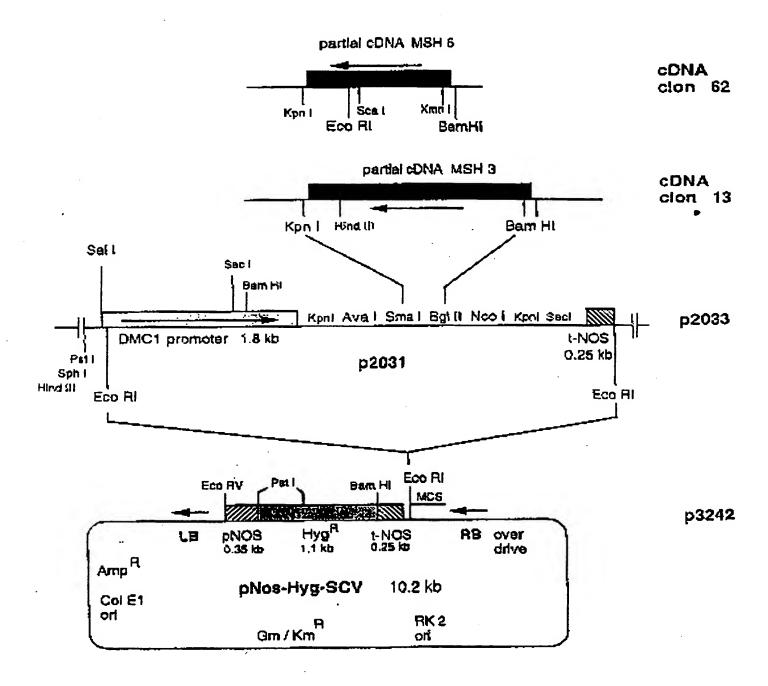
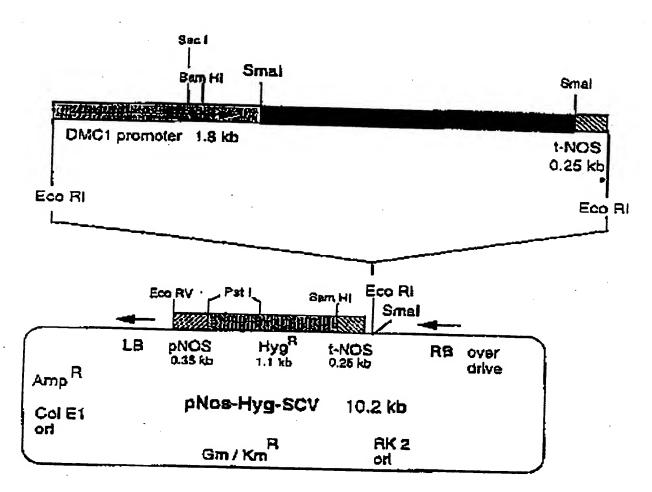


Figure 19

p3243



```
SEQUENCE LISTING
```

```
<110>
               Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
               Marie-Pascale; Freyssinet, Georges; Perez, Pascual.
 <120>
               Methods for obtaining plant variaties
 <130>
               395498C
 <150>
               PO9745
 c151>
               1997-10-10
 <160>
               93
 <210>
               Ł
               33
 <211>
 <212>
               DIVA
 <213>
               Artificial sequence
<220>
 <221>
               modified_base
 <222>
               11
 <223>
 <220>
 <221>
               modified_base
 <222>
               14
 <223>
               I
 <220>
 <221>
               modified base
 <222>
               17
 <223>
               I
 <220>
 <223>
               Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
               ALMSR6.
 c300>
 <301>
               Reenan and Kolodner
 <302>
               Genetica
 <303>
               132
 <306>
               963-973
 <307>
               1992
 <400>
               1
                                                                           23
 ctggatccac nggnccnaay atg
 <210>
 <211>
               23
```

3.

VSDOCID: <WO__9919492A2TI_>

<212>

DNA

```
Artificial sequence
ج 213ء
<220×
<221>
              modified base
<222>
              15
<223>
              I
<220>
<221>
             modified base.
<2223
              19
<223>
              I
<220>
             Degenerate oligonucleotides DOMU used to isolate AtmSH3 and
<223>
             Atmsh6.
<300>
<301>
             Reenan and Kolodner
<3025
             Genetics
<303>
             132
<306>
             963-973
<307×
             1992
<400>
             2
ctggateert artgngthre raa
                                                                        23
<310>
             3
<211>
           24
<212>
             DNA
<213>
             Artificial sequence
<320×
<223>
             MSR3 specific primer 636 for PCR using cDNA of Arabidopsis
             thaliana ecotype Columbia
<400>
             3
tgctagtgcc tcttgcaage tcat
                                                                        24
<210>
             4
<211>
             27
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
             Columbia containing adapter sequences ligated to both its
             enda
<400>
```

ccatcctast	acqueteact atagggc 27
.010	5
<210><211>	33
<212>	DNA
<213>	Artificial sequence
46.60	werrieff podesie
<220>	
<223>	Primer AP2 for PCR using cDNA of Arabidopsis thalians ecotype
	Columbia containing adapter sequences ligated to both its
	ends .
<400>	5
actractata	gggdtcgagc ggc 23
<210>	6
<211>	30
<212>	DNA
<213.	Artificial sequence
	•
<320>	
<2335	MSH3 specific primer S525 for PCR using cDNA of Arabidopsis
	thaliana ecotype Columbia
<400>	б
200110103	tatgtgtgae gettsaetta 30
aggirorgar	01103031345 3510240042
<210>	Ť
<211>	39
<212>	DNA
<213>	Artificial sequence
<220>	
<323≻	MEH3 specific primer S51 for PCR using cDMA of Arabidopsis
	thaliana ecotype Columbia
	_
<400>	7
anat comet a	ctgggttttp agtgtgagg 29
Aberradaara	ctgggttttg agtgtgagg
<210>	
<211>	24
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer 635 for PCR using cDNA of Arabidopais
	thaliana ecotype Columbia

```
<400×
                3
  gracgigett gaiggigtic toac
                                                                          24
  <210>
  <211>
               28
  <212>
               DNA
  <213>
               Artificíal sequence
  <220>
  <223>
               MSH3 specific primer S523 for PCR using cDNA of Arabidopsis
               thaliana ecotype Columbia
  <400>
               9
  tragaragta trragrateg ragaagta
                                                                         28
 <210s
               10
 <211>
               33
 <212>
               DNA
 <213>
              Artificial seguence
 <220>
 <223>
              MSH3 specific primer 1S5 for PCR using cDNA of Arabidopsis
              thaliana ecotype Columbia
 <400>
              10
 atecegggat gggesageas aageageaga ega
                                                                         33
 <210>
              11
 <2115
              27
 <212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             MSH3 specific primer S53 for PCR using cDNA of Arabidopsis
             thaliana ecotype Columbia
<400>
             11
gacaaagagc gaaatgaggc cccttgg
                                                                        27
<210>
             73
<211>
             1250
<212>
             Arabidopsis thaliana acotype Columbia
<213>
<2235
             Clone 52
```

<400> 12

cccgggacgg	дсардсавав	gcagcagacy	acttctcgtt	tettegetes	cdaacccasa	60
terregarte	асдавссдав	teeggtagee	gaateateaa	caccyccacc	gaagatatcc	120
gccactgtat	cettetetee	ttccaagegt	aagottotot	ccgaccatet	càccaccaca	180
teacceasaa	agcetaaaet	ttotoptoae	DOTOSARRO	cagtacccga	tcccaáttta	240
caccaaagat	ctcccagag	atttctggsa	ccctcgccgg	aggaatatgt	tecegaaacg	300
ccatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattettegg	sgaagacgcg	420
gagatogcag	cacgegtgtt	gggtatttac	gctcatatgg	atcaceattt	cstgscggcg	480
agtgtgccaa	cattlegate	gaausuccai	gtgagaagac	tggtgaatgc	aggatacaag	540
atcggtgcag	cgaagcagac	tgaaactgca	gccartsagt	cacatggtge	аласоддаес	600
ggccottttt	tcoggggact	gtcggpgLlg	tataccasag	ccacgcttga	ageggetgag	660
gatataagtg	gragregtag	tggtgaagaa	ggttttggtt	cacagagtaa	tttettggtt	720
tgtgt tg t gg	atgagagagt	taagtoggag	acattagget	gtggtattga	aatgagtttt	780
gatgttagag	toggtgttgt	tggegttgaa	atttcgacag	gtgaagetgt	tcatgaageg	840
ttcaatgata	actteatgag	aagtggatta	gaggetgtga	ttttgagett	gtcaccagct	900
gagctgttgc	ctggecagec	totttcacaa	c asac tgaga	agtt.tttg gt	gycacatyct	960
ggacctarct	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	csacggtaat	1020
gcagtagatg	aggttattte	attatgtgaa	aasatcagcg	caggtaactt	agaagatgat	1080
aaagaaat ga	aŭttādaâdc	tgctgaaaa	ggsatgtett	gcttgacagt	tcatacaatt	1140
atgaacatge	cacatrigae,	tgttcaagrc	ctegecetaa	cgttttgcca	tctcaaacag	1200
tttggatttg	aaaggateet	ttaccaaggg	gestsættte	getetttgte		1250

<210> 13 <211> 34 <212> DNA <213> Artificial sequence

<220>
<223> MSH3 specific primer 255 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400>	13					
atconggys	C aaaatgaac	agreggetti	agto			34
<210ء	14					
<211>	27					
<212>	ONY					
<213>	Artifici	sodneuce	:			
<220>			`			
<223>	MSH3 spe	cific brime	r csa for c	00B		
	chaliana	ecotype Co	lumbia	ck using C	ONA of Arabid	eisgo
<4DO>	14					
gccacatete	actgttcasq	ccctcgc				27
						٤r
<210>	15					
<211>	2110					
<212>	ANG					
<213>	Arabidop	aniind aiz	a ecotyme C	in larmon is		
<2235	Clone 13					
<400>	1.5					
gccacatetg	actgttcaag	cectegeect	aacgttttgc	catotosaac	agtttggatt	50
tgaaaggatc	ctttaccaag	gggcctcatt	togetetttg	tcaagtdaca	cagagatgac	120
CEEEEcagec	autactotgo	aacagttgga	ggttgtgaaa	aataatteag	atggatcgga	180
atctggetce	ttattecata	atatgaatca	cacacttaca	gratatggtt	ccaggettet	240
tagacactgg	gugactcate	ctotatgcga	tagaaatttg	statetgeto	ggcttgatgc	300
tgtttctgag	atttetgett	gcatgggatc	teatagttet	toccagetca	gcagtgagtt	360
ggttgaagaa	ggttctgaga	gagcaattgt	ateacetgag	tretatoreg	tgctctcctc	420
agtettgaca	gctatgtcta	gatcatctga	tattcaacgt	ggaataacaa	gaatetttea	480
toggactget	asagccacag	agtteattge	agttatggaa	gctattttac	tigoggggaa	540
gcaaattrag	cggcttggca	tasagcasga	ctctgaaatg	aggagtatgd	astetgease	600
					ttgtggttga	660
					gtgacttget	720
		gcgaccaatt				780
	`				agctcgctat	840

togaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
trocaaggto	cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	ttegatatea	960
cocccagaa	acageagetg	gcttggatga	gctagcceta	gcsactgaar	arcttgccat	1020
tgtgaaccga	åcttcätäåä	atagtttect	caagagtttc	agcagatact	acacagattt	1080
taaggetgee	gtteaagete	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacsag	aactatgtcc	gtecegagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtct	ggtrgtcate	ctgtactgga	gactatatta	caagataact	tegteccaaa	1260
tgacacaatt	regeatgeag	aaggggaata	ttgccaaatt	atcaccggae	ctaacatggg	1320
aggaæagagc	tgetatatee	gtcaagttgc	tttaatttcc	ataatggctc	aggttggttc	1380
otttgtacca	gcgtcattcg	ccsagetgea	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgetteagae	agtatccagc	atggcagaag	tacetticta	gaagaattaa	gtgaagegte	1500
agaqataatq	agaacntgit	cttctcgttc	gettgtesta	ttagacgagc	trggaagagg),560
cactagcaca	<u> </u>	sagocasugo	ctatgcaaca	ttacagoato	toctageaga	1620
aaagagatgt	ttggttcttt	ttgtczcgca	ttaccctgaa	atagetgaga	tcagtaacgg	1690
attoccaggt	setgttggga	cataccatgt	ctegtatetg	acattgcaga	aggataaagg	1740
cagtratgat	catgatgatg	egacctacet	alutaagcti	gtgcgtggtc	tttgcagcag	-1900
gagetttggt	tttaaggttg	oteagettge	ccagatacct	ccatcatgta	tacgtcgage	1860
catttcaatg	gctgcaaaat	tggaagetga	ggtacgtgca	agagagagaa	atacacgcat	1920
ддадаасса	gaaggacatg	aagaaccgag	aggegeagaa	gaatctattt	cggctctagg	1980
tgacttgttt	gcagacctga	aatttgctct	ctctgaagag	gacccttgga	aagcattcga	2040
gttttaaag	catgettgga	agattgctgg	caaaatçaga	ctaaaaccaa	cttgttestt	2100
ttgacccggg						2110

<210>	16
<211>	29
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSK3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	16

3	gatc	à à à g	a ct	aggt!	tttg	agt	gtga	33										29
<:	210 ₂		j	17														
٠.	711>		3	9 0														
<:	212>		r	NA														
<2	213>		Д	Artif	icia	l se	quer	ice										
< 2	20>																	
	223>		2.6	(CH 2		2 = 2 .												
			t	hali	ana	6505	Abe brī	coln	9525 mbia	for	PCR	у паз	ing (ONA	ot A	rabid	lopai	8
< 4	00>			7														
зg	gttc	tgat	tat	gtgt	gac	gctt	tact	ta										30
<2	10>		1	8														
د 2	11>		3	522														
< 2	12>		Ð	NA														
< 2	13>		A:	rabi	ತೆಂಬಕ.	رج خا	taii.	×122 4	rcot;	Abc	Colu	mia					•	•
<2	20×																	
<2	21>	•	CI)S											,			
	22>		()	100).	1	(3342	()											
€2	23>		At	:NS)(2	fu)	1-16	ngti	CON	is Af	ಾರೆ ನೇ	educe	م امد	.		. 4 _ \.	e en		_
			Pc	y Abe	ptid	ie							2 d a C 1	106 (31. GI	e en	<u> ೧೮೮೯</u> ೦	l
<40) 0>		18	:														
ccs	aage	asag	c ಡಿ ಇಡಿ	rcgaa	aa t	.tggc	aacc	c as	gtto	geca	ı tag	3CCa	gae	CACC	;acce	cc.	60	ı
					ga g											- 0		
								•									99	
Met	Glv	Lva	Gla	Lve	Cag	csá	acg	atc	tot	cgt	ttc	ttc	: gct	ccc	aaa		147	
1		-,0	U 1.77	5 5	GIN	GIN	Inr	Ile	ser	Arg	. 5pe	Phe	: Ala	Pro	tys Lys			
				•					10					15				
CCC	aaa	tcc	ccâ	act	CBC	gaa	cea	aat	CCD	nt a	~~~				aca			
510	Lys	Ser	Pro	Thr	His	Glu	Pro	Asn	Pro	Val	2) 9	GJ/I	COR	CCa	aca Thr		195	
			20					25			7.2.0	0.4	30	ser	Thr			
ccg	CCE	crg	aag	ata	tee	ade	acr	ôt-s	*	** -						`		
Pro	Pro	Pro	Lys	Ile	Ser	Ala	Thr	Maj Are	Can	Bho	COL	CCE	TOO	aag	cgt		243	
		35				- 10 01	40	7 10 3	562	FIIC	eci.	9F0 45	ser	Lye	Arg			
aag	ctt	cte	tee		-114		.											
Lys	Leu	Leu	Ser	Ago	cac	Tan	gcc	gcc	āca	tca	ccc	238	23 g	CCL	aaa		291	
-	50		••		His	55	wig	HTS	ALA	ser	Pro 60	Lys	Lys	Pro	Lya			
											_							
CCE	CCE	CCE	Cac	act	caa	aac	CCA	gta	ccc	gat	ccc	aat	tta	cac	CZZ		339	
65	SEL	PIO	HIB	Thr	944	Asn	510	Val	Pro	Asp	Pro	Asn	Leu	Kis	Gln		200	
45					70					75					80			

aga Arg	ttt Phe	Den Len	eyv cad	aga Arg 35	ttt Phe	ctg Leu	ga∂ Glu	51.0 GGG	Ser 90	ceg Pro	gag Glu	gaa Glu	tat Tyr	gtt Val 95	ece Pro	387
gaa Glu	acg Thr	t.ca Ser	cca Ser 100	tcg Ser	agg Ar g	ada Lys	tac Tyr	aca Thr 105	cca Pro	t t g Leu	gaz Glu	cag Gln	caa Gln 110	gtg Val	gtg Val	435
gag Glu	cta Leu	aag Lys 115	agc Ser	Lys	tac Tyr	Pro	gat Asp 120	gtg Val	gtt Val	ttg Leu	atg Met	9tg Val 125	gaz Glu	gtt Val	ggt Gly	483
tac Tyr	Arg Arg	tac Tyr	aga Arg	ttc Phe	trc Phe	934 Gly 135	gaa Glu	Asb Gsc	A)a gcg	gag Glu	atc Ile 140	gca Ala	gca Ala	Arg	gtg Val	531
ttg Leu 145	GIY ggt	att	tac Tyr	gct Ala	cat His 150	atg Met	gat Asp	cac His	аас Авл	ttc Phe 155	atg Met	acg Thr	VJ9 3cd	agt Ser	160 gtg	579
cca Cca	aca Thr	ttt Phe	Arg cga	ctg Low 165	aac Asn	ttc Phe	cat His	gtg Val	aga Arg 1.70	aga Arg	ctg	gtg Val	aat Asn	gca Ala 175	gga Gly	627 ₽
						aag Lys										67S
cat Bis	g]y ggt	gca Ala 195	aac Asn	Ar3 cgg	acc Thr	C1 y ggc	CCT Pro 200	uut Phe	ttc Phe	yığ cg3	ejy èga	ctg Len 205	tcg Ser	gcg Ala	t Eg Leu	723
tat Tyr	acc Thr 210	aaa Lys	gcc Ala	acg Thr	ctt Leu	gaa Glu 215	gug Ala	gct Ala	G); G);	gat Asp	ata 1]e 220	agt Ser	ggt Gly	gyy ggt	tgt Cys	771
99t Gly 225	gly ggr	gaa Glu	gaa Glu	ggt Oly	Phe 230	ggt Gly	tca Ser	cag Gln	agt Ser	aat Asn 235	tto Phe	ttg Leu	gtt Val	tgt Cya	gtt Val 240	819
gtg Val	gat Aap	gag Glu	aga Arg	gtt Val 245	aag Lys	tog Ser	gag Glu	aca Thr	tta Leu 250	gg¢ gg¢	tg t Cys	gly ggt	att Ile	gaa Glu 255	atg Met	867
agt Ser	ttt Phe	gat Asp	gtt Val 260	aga aga	gtc Val	ggt	grt Val	gtt Val 265	ggc Gly	gtt Val	gaa Glu	att Ile	ccg Ser 270	aca Thr	ggt ggt	915
gaa Glu	gtt Val	gtt Val 275	tat Tyr	gaa Glu	gag Glu	ttc Phe	aat Aen 280	gat Asp	Jas α≎A	ttc Pha	atg Met	aga Arg 285	agt S er	gga Gly	tta Leu	963

									10							•
G 1 35	19 gc .u Al 25	stg la Va ao	tg a al r	tt t: le Ge	eu Sei	C ttg Leu 295	Set	Pro	a ge	t gaq a Glo	g cti U Lei 300	r La	g ct u Le	t gg	c cag	1011
30	5			•	310)	БУБ	PNE	. rer) Val	L ALE	3 Met	t Al	a Gl	a cct y Pro 320	i059
				32	5		n1 g	MIG	330	. Ten	l Asp	cys	s Pho	€ Se: 33!		1107
			34	0	t gag p Glu	. val	116	345	ren	che	Glu	Lys	350	e Sei	r Ala	1155
		35	5	- 7.0	c gat p Asp	Lys	360	(aG.E	Lys	Lou	Clu	Ala 365	Ala	t Glu	Lyb	1203
-	370	;	,	J 22.	g aca 1 Thr	375	n.u.	ınr	TIC	Mct	A513 380	Met.	Pro	Ris	Leu,	1251
385				5 500	390 390	LOU	inr	rne	Cys	815 395	Геп	Lys	6),n	Phe	Gly 400	1299
		_	,	405		CIN	ota.	Ala	410	Phe	Arg	Ser	Leu	Ser 415	Ser	1347
			420)	cto Leu	ser ,	n_a .	425	Thr	Leu	Gln	Gln	10u 430	gag Glu	gtt Vaļ	1395
_	-,-	435	ven	ser	gat Asp	GIA 8	er (31 u	Ser (Gly .	Ser	Leu 445	Phe			1443
	450			20311		155	yr u	-1Y (ser)	arg i	Leu 1 460	Leu	yrg	Hie	Trp	1492
£65				Dea	tgc g Cys 2 470	eb Y	ig A	isn i	Leu 1	(le 3 175	Ser A	Ala.	Arg	Геп	Aap 460	1539
got, Ala. S	gtt Val	tct Ser	gag Glu	att Ile 485	tet 9 Ser A	ict to	gc a ys M	et (gga t Sly s	er H	at <i>d</i> Rie S	er s	3er	tcc Ser (cag Gin	1587

WO 99/19492 PCT/EP98/06977

								ggz ggz								1635
cet Pro	gag Glu	ttt Phe 515	tat Tyr	CEC Leu	gtg Val	ctc Leu	tcc Ser 520	tca Ser	gtc Val	ttg Leu	aca Thr	gct Ala 525	atg Met	tct S er	aga Arg	1633
								aça Thr								1731
								atg Met								1779
								aag Lys								1827
								act Thr 585								1875
								yab								1923
								cga Arg								1971
			-	-				gag Glu		_	_	_	_		_	2019
								gat Asp								2067
aag Lys	aag Lys	ctc Leu	gct Ala 660	att 11e	cga Arg	aat Asn	ttg Leu	gaa Glu 665	ett Phe	ctt Leu	caa Gln	gtg Val	teg Ser 670	617 833	atc Ils	2115
								gar Asp								2163
								act Thr								2211

										• -							
70	5					71	.0	LQ 12.	ın Mi	ia Le	14 DE	.a. Th .5	r GJ	u Ki	s le	ε τ gc 24 Αὶ 72	a 0
				-	72	5		PAC	,D 26	73	о :s те	u Ly	5 Se:	r Ph	e Se 73	_	9
	•			740		,	~ ~1	a vi	74	1 G1: 5	n Al	a Le	u Ala	A1 75	a Le O	n ysl	P
·		7	55			, U.C.		76	0	r Ar	g asi	n Ly:	769	Ty	r Va	c egt l Arg	3
	77	0			, w p	V1	77:	5	t ere	o val	l Gli	780	? Авл >	110	e Gl:	g tot n Ser	
990	. cg	ت ر - ا	at e	CCE	gr.a	cro	ga	3 act	اغلا	ı tta	cae	ya;	aac	ttr	. ara	cca	2499
785	•	-			,	790) U1.	.6 .2	- 116	2 TSA	795	y Ysp	ysn)	Pla€	· Va	Pro 808	•
				- -	805	20		`	I GIY	810	YY	Cys	Gln	Ile	915 915		
•			8	20	o Ly	917	гуs	96L	625	Tyr	Ile	Arg	Gln	Val	νjs	tta Leu	2595
tts ell	Ser	at 11 83		tg d	gct Ala	cag Gln	gtt Val	ggc Gly B40	zer	t:: Phe	gca Val	cca Pro	grg Ala 845	tca Ser	ttc Phe	Pla	2643
rya	ctg Leu 850	Ca Hi	s V	tg d al 1	cet . Leu	gat Asp	855 617 99t	gtt Val	ttc Phe	açt Thr	yrg ¢gg	atg Met 860	ggt Gly	gct Ala	tca Ser	gac Asp	2691
agt	atc	Ca	ġ ca	at 5	ge	aga	agt	acc	ttt	cta	gaa	gaæ	tta	agt	naa	aca	2739
965					,	970	≎e1.	MA	Pne	Leu	611 875	G1u	Leu	Ser	Glu	Ala 880	4739
tca Ser			:	8	85		che	SEL	ser	890	ser	Leu	Val :	Ila	Leu 895	Asp	2787
gag Glu	ctt Leu	G1A Gas	25 / Ax 90	-	ly 1	het i	agc Ser	1117	cac Kis 905	gac Aap	ggt	gta (Val)	Ala 1	ile Blo	gcc Ala	tat Tyr	2835

WO 99/19492

gca Ala	aca Thr	eta Leu 915	çağ Gln	cat His	CEC Lou	cta Leu	gca Ala 920	gaa Glu	aag Lys	aga Arg	cgt Cys	ttg L o u 925	gtt Val	ctt Leu	ttt Phe	2883
														cca Pro		2931
														gat Asp		2979
														gtg Val 975		3027
														gcc Ala	_	3075
						Arg					Mec			aaa Lys		3123
Glu	_		-		Ala					Thr		-		gaa Glu		3171
_	gly		-	Glu					Glu				_	gct Ala		3219
			Spe	_				Phé	-				${\tt Gl} u$	gac Asp L055		3267
		Ala					Lув		-		ľ'ns	Ile	-	GJA aac		3315
	Arg	cta Leu 1075				Cys			tgai	ttaa	atc 1		catta	3 E		3362
∌åcs	acto	gca a	iggt (ttga	it ca	stetg	gttag	g ttg	gegti	ecta	acti	atgi	tgt a	etta:	gtataa	3432
caaq	jaaaa	aga (jaati	agag	ga ga	atgga	kttc1	t aat	ccdi	gtgt	tgca	agtad	cat (ettt	tetèca	3482
cccó	;cata	aaa a	aaaa	aaa	la a	aaaa	aaaa	aaa	aaaa	asee					-1	3522

۷.	213.	,			1002	dops	is t	hali	222			-	mbia			
< 3	323:	•		1	Polyp	epri	රාද M	12H3	ana	6000	.ype	Colu	mbia			
	000				19											
										,	, Q				15	
									2	5				3 (•	Thr
			-	•				4.	0				45	}		Arg
							3	>				60)			Lye
						7.0	,				79	5			His	80
					3.	,				.90	•				. Val 95	
				101	•				105	•				110		
								120	1				125		Val	
		_					135			•		340			Arg	
						130					155				Ser	160
					100					170					Ala 175	_
Tyr	Ly	\$ }	Ele _.	Gly 180	Val	Val	Lys	Gln	Thr 185	Glu	Thr	Ala	Ala	Ile 190	Lys	Ser
His	G1;	7 7	la .95	asa	Arg	Thr	Gly	Pro 200	Phe	Phe	Arg	Gly	Leu 205	Ser	Ala	Lev
Tyr	Th:	: 1	ys	Ala	Thr	Leu	GԼս 215	Ala	Ala	Glu	Asp	11e 220	Ser	Gly	Gly	Cys
Gly 225	Gly	G	lu	Glu	Gly	Phe 230	G1y	Ser	Gln	5er	Asn 235	Phe	Leu '	Val (Cys :	Val 240
Val	Asp	G	lu	Arg	Val 245	LyL	Ser	Glu	Thr	Leu 250	Gly	Сув	GJY :		31u 355	Ket

ser	Phe	Asp	360 Val	Arg	Val	Gly	Val	Val 265	Gly	Val	Glu	Ile	5er 270	Thr	Gly
Gle	Val	Val 275	Τγε	glu	Glu	Phe	Asn 230	Asp	Asn	Phe	Met	Arg 285	Ser	GJY	Leu
ĠÌu	Ala 290	Val	Ile	Leu	Sor	Leu 295	ser	Pro	Ala	Glu	300 300	Leu	Lev	Gły	nte
205	Гел	3er	G1u	Gln	Thr 310	Glч	Lys	Phe	Leu	Val 315	Ala	Met	Ala	Gly	Pro 320
Thr	Sex	ΩeA	Val	Arg 325	Val	Glu	yrg	Ala	8er 330	ьеч	Yeb	Суа	Phe	Ser 335	Aen
Gly	Asn	Ala	Val 340	Asp	Ġlи	Val	lle	Ser 345	Leu	Суз	Glu	ГÀЗ	11e 350	ser	a <i>l</i> a
Gly	Asn	1.∉n 355	Glu	Asp	Asp	Lys	G1u 3&0	Met	Lys	Leu	Сĵи	Ala 385	Ala	Glu	ГÀВ
Cly	Met 370	Ser	Cys	ren	Thr	Val 375	Eis	Tinr	Lle	Met	300 ash	Met	Bra	His	Leu
Thr 385	Val	Gln	Ala	ľėn	Ala 390	Lcu	Thr	Phe	САз	llis 395	Leυ	Lys	Gln	Phe	Gly 40 0
Phe	Glu	Arg	Ile	1.eu 405	Tyr	Gln	ĠĴУ	Ala	Ser 410	Phe	Arg	Sex	Leu	Ser 415	Ser
Asn	Thr	Glu	Met 420	Thr	Leu	Ser	Άla	Asn 425	Thr	Leu	Gju	Gln	Leu 430	₿}u	Val
Val	Lys	Дви 435	σ8 Α	Ser	дæҚ	Gly	3er 440	Glu	Бer	ፅያላ	Ser	Leu 445	Phe	His	ask
Met	Asn 450	His	Thr	Leu	Thr	Val 455	Tyr	Gly	Ser	Arg	Leu 460		Arg	ਮ3.¢	Trp
Val 465	Thr	His	Pro	Leu	Суя 470	Asp	Arg	ABD	Leu	11e 475		Ala	Arg	Leu	ASP 480
Ala	Val	Ser	Glu	1le 485	Şer	Ala	Суз	Met	Gly 490		eiK	Ser	Ser	Ser 495	Gln
Leu	Ser	Ser	Glu soo	Leu	Val	Glu	Çlu	G1y 505		Glu	Arg	Ala	Ile 510	Va1	ser
Pro	GŢņ	9he 515	Tyr	Гел	Val	Leu	Ser 520		Val	Leu	Thr	Ala 525		Şer	Ärg
Ser	Ser 530	_	Ile	Gln	_	Gly 535		Thr	Arg	11e	Pha 540		Arg	Thr	Ala

- Lys Ala Thr Glu Pho Ile Ala Val Met Glu Ala Ile Leu Leu Ala Gly
 545 550 555 560
- Lys Glo Ile Glm Arg Leu Gly Ile Lys Gln Asp Ser Glu Met Arg Ser 570 575
- Met Gln Ser Ala Thr Val Arg Ser Thr Leu Leu Arg Lys Leu Ile Ser 580 590
- Val 11e Ser Ser Pro Val Val Val Asp Asn Ala Gly Lys Leu Leu Ser 595 600 605
- Ala Leu Asn Lys Glu Ala Ala Val Arg Gly Asp Leu Lau Asp Ila Leu 610 620
- Ile Thr Ser Ser Asp Gln Phe Pro Glu Leu Ala Glu Ala Arg Gln Ala 625 630 635 640
- Val Leu Val Ile Arg Glu Lys Leu Asp Ser Ser Ile Ala Ser Phe Arg 645 650 655
- Lys Lys Leu Ala Ile Arg Asn Leu Glu Phe Leu Gln Val Ser Gly Ile 660 665 670
- The His Leu Fle Glu Leu Pro Val Asp Ser Lys Val Pro His Asm Trp 675 680 685
- Val Lys Val Asn Ser Thr Lys Lys Thr Ile Arg Tyr His Pro Pro Glu 690 695 700
- The Val Ala Gly Leu Asp Glu Leu Ala Leu Ala Thr Glu His Leu Ala
 705 710 715 720
- Ile Val Asn Arg Ala Ser Trp Asp Ser Phe Lou Lys Ser Phe Ser Arg
 725 730 735
- Tyr Tyr Thr Asp Phe Lys Ala Ala Val Gln Ala Leu Ala Ala Leu Asp
 740 745 750
- Cys Leu His Ser Leu Ser Thr Leu Ser Arg Asn Lys Asn Tyr Val Arg 755 760 765
- Pro Glu Phe Val Asp Asp Cys Glu Pro Val Glu Ile Asn Ile Gln Ser 770 780
- Gly Arg His Pro Val Leu Glu Thr Ile Leu Glu Asp Asn Phe Val Pro
 785 790 795 800
- Asn Asp Thr Ile Lau His Ala Glu Gly Glu Tyr Cys Gln Ile Ile Thr 805 810 815
- Gly Pro Asn Met Gly Gly Lys Se: Cya Tyr Ile Arg Gln Val Ala Leu 820 825 830

St. Santa Land

Ile	Ser	I le	Met	Ala	Gln	Val	Gly	Ser	Phe	Val	Pro	Ala	Ser	6)ಗಳ	Ala
		835					849					845			

- Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp 850 855 860
- Ser Ile Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala 865 870 875 380
- Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp 885 890 895
- Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr 900 905 910
- Ala Thr Leu Glo His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe 915 920 925
- val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Asn Gly Phe Pro Gly 930 935 940
- Ser Val Gly Thr Tyr His Val Ser Tyr Leu Thr Leu Gin Lya Asp Lyc 945 950 955 960
- Gly Ser Tyr Asp Bis Asp Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg 965 970 975
- Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln 980 985 990
- Ile Pro Pro Ser Cys Ile Arg Arg Ala Ile Ser Met Ala Ala Lys Leu 595 1000 1005
- Glu Ala Glu Val Arg Ala Arg Glu Arg Ash Thr Arg Met Gly Glu Pro 1010 1015 1020
- Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser lle Ser Ala Leu 1025 1030 1035 1040
- Gly Asp Leu Phe Ala Asp Leu Lye Phe Ala Leu Ser Glu Glu Asp Pro 1045 1050 1055
- Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys
 1060 1065 1070

⊁

11s Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080

<210> 20

<211> 24

<212> DNA

<213> Artificial sequence

<230>		
<223>	MSB6 specific primer 638 for PCR using cDNA of Arabidops thaliana acotypa Columbia	sis
<400>	. 20	
cototaco	ond Bedgedgess seed	24
<210>	21	
<211>	28	
<212>	DNA	
<2135	Artificial sequence	
<220>		
<223>	Primer 381 for PCR using cDNA of Arabidopsis thaliana ec Columbia	otype
<400>	21	
cgtcgcctt	tt aggatecoot tootreae	28
<210>	23	
<211>	30	
<212><213>	DNA	
(423)	Artificial sequence	
<220>		
<223>	MSH6 specific primer 5823 (or PCR using cDNA of Arabidop	
	thaliana ecotype Columbia	615
<400>	22	
gcttggcgc	ca telaatagaa teatgaragg	30
	-	
<210>	23	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MEUS specific primer (17 for any	
	MSR6 specific primer 637 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	is
400>	23	
] ತಿ ದಕಿಡೆಂಡೆದಳಿ	a gttottcaga atgo	24
	· ·	
210>	24	
211>	33	
212>	DNA	

WO 99/19492 PCT/EP98/06977

<213>	Artificial sequence	
<230×		
	Atolic	
4223×	MSH6 specific primer 188 for FCR using cDNA of Arabidopsi thaliana ecotype Columbia	İs
<400×	24	•
ateregggat	gcagcgreag agategailt tgt	33
214	· ·	
<210>	25	
<211>	27	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer SB3 for PCR using cDNA of Arabidopsi	
	thaliana ecotype Columbia	
<400>	`25	
	•	
ogotatetat	ggetgetteg aattgag	27
<210>	26	
<211>	1385	
<211>	CNA	
<213>	Arsbidopsis thaliana ecotype Columbia	
<223>	Clone 43	
<400>	26	
cccgggatgc	agogocagag atogattitg tottletter aaasareese ggeggegact	60
argaagggtt	tggtttccgg cpatgctqct agcggcgggg gcggcagcgg aggaccacga	120
tttäätgtga	aggaagggga tyctaaaggc gacgcttctg tacgttttgc tgtttcgaaa	180
tetgtegatg	aggitagagg aacggatact ccaccggaga aggitecgcg tcgtgtcctg	240
ccgtctggat	ttaagorggr tgaatorgre ggtgatgett egteeetgtt etecaatatt	300
atgcataagt	ttgtaaaagt cgatgatcga gattgttctg gagagaggag ccgagaagac	360
gttgttccgc	tgaatgattc atctctatgt atgaaggcca atgatgttat tcctcaattt	420
cgttccaata	atggtaaaac tcaagaaaga aaccatgett ttagttteag tgggagaget	480
gascttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
gggatgcgtc	cacqtqcttc tcgcttgaag cgaqttctgg aggatgaaat gacttttaag	600
gaggaraagg	tecetgtatt ggactetaac amanggetga aaatgeteea ggateeggtt	660

cgeggagag:	a sassagaag	t aaacgaagg	a accadattt	g aatggcetg	a gtottotoga	720
atcagggat	g ccaatagang	g acgtectga	t gateeretti	t acgatagaa	a gaccttacac	780
ataccaects	3 atgicticae	a gaaaatgte	o geateacaaa	agcaatatt	g gagtgttaag	840
agtgaatat	tggacattgt	gettttett	t a as gtgggg	a aattttatg	gctgtatgag	900
ctagatgcg(y aattaggte:	ı caaggageti	gactggaaga	tgaccatgag	t tggtgtggga	960
asatgeagad	aggttggtat	: ctctgaaagt	gggatagatç	3 aggcagtgca	a aaagctatta	1020
gctcgtggat	: atasagttgg	acgaatcgad	i cagotagaaa	catetgaeca	ı ağcaasagco	1080
agaggtgcta	atactataat	cccaaggaag	ctagttcagg	tattaacte	: atcaacagca	1140
9Acd9Add8	acatogggco	tgatgccqtd	: catettettg	r ctacaaaaga	gatcaaaatg	1200
gagctacasa	agtgttcaac	tgtgtatg g s	tttgettttg	ttgactgtgc	: tgccttgagg	1260
ttttgggttg	gyterateag	ျားလူကရောက်ရောက	tratgtgetg	cccrrggago	gttattgatg	1320
caggettete	caaaggaagt	gttatatgac	agcaaagggc	tatesagaga	agcacaaaag	1330
getetaagga	aatatacgtt	dacsäddrot	acggrggcac	agttggctcc	: agtaccacaa	1440
gtaatggggg	atacagatgo	tgctggagtt	agaaatataa	tagaatotaa	cggatacttt	1500
aaaggetett	cegaatcatg	gaactigcgct	gttgatggtc	taaatgaatg	tgstgttgcc	1550
cttagtgete	ttggagaget	asttaatcat	ctgtctaggc	taaagctaga	agatgtactt	1620
aagcatgggg	atatttttcc	ataccaagtt	tacaggggtt	geoteagaat	cgatggccag	1580
acgatggtaa	atcttgagat	atttaacaac	agctgtgatg	gtggtecsta	agggacctrg	1740
taceaatate	ttgatametg	tgttagteca	actggtaage	gactottagg	gaattggatc	1800
tgccatccac	tessagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaatte	1860
acggcaaact	cagaaagtat	ġca44tcact	ggccagtate	tocacaaact	tocagactta	1920
gaaagactgc	toggacgcat	caagtotago	gttogatoat	cagostotgt	gttgcctgct	1980
cttctgggga	aaaaagtgct	gaaacaacga	gtraaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatcc	gttgrrggct	ctacagaagg	aaccaaacat	gatgagtttg	2100
ctttataaac	totgtasact	toctacatta	gtaggaaaaa	gcgggctaga	gttatttctt	2160
tetcaatteg	aagcagccat	agatagçç				2195

WO 99/19492

1250

1320

1390

<210>	27					
<212>	1385					
<212>	DNA					
<213>		sia ebalia				
		sie thaliana	a ecotype co	בוטחשוני		
<223>	Clone 62					
<400>	27					
catcageete	tgtgttgcet	getettetgg	ggaaaaaagt	getgaaacaa	cgagttamag	60
catttgggca	aattgtgaaa	gggttcagaa	gtggaattga	tetgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	aactctgtaa	acttcctata	ttagtaggaa	130
aaageggget	agagttattt	ctttctcaat	tegnageage	catagatage	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgetgaaac	teteaesata	cttatçgaac	300
tttttatega	aagagcaact	caatggtetg	aggtcattca	Caccataagc	ogectagatg	360
tortgagate	ctuegoaarn	gcagcaagtc	tototgetgg	aagcatggcc	aggeetgtra	421
				поваддеска		480
					gttccgaatg	. 540
•				rearceregg		600
					ctggccgtta	, 660
				cgaaatotcc		720
				aggagagagt,		784
				tcaggattca		84(
				tgccattgca		900
					taccaccete	
					tgcgcattca	
					ttgtaccgtt	
					gctggaatac	
caaaccaagt	ggttgaaaca	gcatcaggtg	ctgctcaagc	catgaagaga	tcaattgggg	1200

addacticad greadgigag claddatetg agrictedag tergeargad gaerggerea

agreattggt gggtatttct cgagtcgccc acaacaatgc ccccattggc gaagatgact

acgacacttt gttttgctta tggcatgaga tcaaatcctc ttactgtgtt cccaaataac

```
ccggg
                                                                         1385
 <210×
               28
 <211>
               34
 <212>
              DWA
 c213>
              Artificial sequence
 <220>
              MSH6 specific primer 288 for PCR using cDNA of Arabidopsis
 <223>
              thaliana ecotype Columbia
 <400>
              26
 atorogggtt attigggman anagtaagag gatt
                                                                          34
 <210>
              29
 <211>
              27
 <212>
              DNA
 <213>
              Artificial sequence
 <220×
             MSH6 specific primer SB2 for PCR using cDNA of Arabidopsis
 <233>
              thaliana ecotype Columbia
<400>
              29
gegttegate atcagectet gtgttg:
                                                                          27
<210>
             30
<211>
             3606
<212>
             Arabidopsis thalisms ecotype Columbia
<213>
<220>
<221>
             CDS
<222>
             (142) . . . . [3468]
             AtMSH6 full-length command deduced sequence of the encoded
<223>
             polypeptide
<400>
             30
aaaagttgag ccctgaggag tatcgtttcc gccatttcta cgacgcaagg cgaaaatttt
                                                                        60 .
tggcgccaat ctttcccccc tttcgaatte tctcagctea aaacatcgtt teteteteae
                                                                       120
teteteteae aattecaasa a atg cag ege cag aga teg att tig tet tie
                                                                       171
                        Met Gln Arg Gln Arg Ser Ile Leu Ser Phe
```

								acg Thr								373
				15					50					25		
get	gct	agc	33c	999	ggc	ggc	agc	gga	gga	cca	ega	ttt	ast	gtg	aag	257
s1A	Ala	ser	Gly	Gly	GJA	CJA	Ser	GTA	Gly	Pro	Arg	Phe	Asn	Val	Arg	
			30					35					40			
								tct Ser								315
Ģ1u	GIY	45	AIG	Dy 6	31y	~sp	50	361	ANI	urd	rn e	55	447	⊅4T.	гув	
		-					-	gat					_	-	_	363
Ser	Val 60	Asp	GLU	VAI	arg	65 65	Thr	Asp	Thr	Pro	70	GT#	Lya	Val	Pro	
											•					
								aag Lys								411
75	wrd	Val	Den	PLO	30	BIA	51:6	гув	PIO	85	σıα	Se f.	мта	GIA	90 80	
									•							
			_					atg				-			_	459
ALA	Sex	ser	гел	Ph@ 95	ser	ASS	718	Men	1115	ΓĀΡ	hve	Val	Lys	105	Asp	
				73					100					103		
_								agc								507
Asp	уzā	Aep	_	Ser	Gly	Glu	ул.ä		Arg	Glu	Asp	Val		SLO	ren	
			110					115			•		120			
								gct								555
U\$¥	yeb		Ser	Гел	cys	Met		Ala	aea	Asp	Val		Pro	Gln	Phe	
		125					130					135				
								gaa	_			-		-		603
Arg		Asn	ne4	GJY	Lys		Sin	รใน	Arg	Aen		Ala	Phe	ser	Phe	•
	140					145					150					
_		-	_	•		_		gta	_	_				_		651
	Gly	Arg	Ala	Glu		Arg	Ser	Val	Glu	-	ITS	GTA	Val	увр		
155					160					165		•		•	170	•
gat	gtt	cct	ggt	cca	gaa	aça	cca	999	atg	cgt	cca	cgt	305	tct	cgc	69 9
Aep	Val	Pro	gjå		Glu	Thr	Pro	Gly		Arg	Pro	Arg	Ala		Arg	
				175					780					135		
ţţģ	aag	cga	gtt	ctg	gag	gat	gaa	atg	act	ttt	aag	gag	gat	aag	gtt	747
ยอน	Lys	Arg		Leu	Glu	Asp	Glu	Met	Thr	bµe	Ļув	Glu	Asp	Lys	Val	
			190					195					200			
ççt	gta	ctg	gac	tet	aac	aaa	agg	ctg	228	arg	cte	cag	gat	ccg	gtt	795
		Leu						Leu				Gln				
		205					310					215				

CÀ8 FâF	gga Gly 220	Giu	Lys dag	ŗķe	gaa Glu	gta Val 225	aac Asn	gaa Glu	gga Gly	acc	аза Lys 230	ttt Phe	gaa Glu	tgg Trþ	ott Leu	943
gag Glu 23 5	ser	ts: Ser	cga Arg	ace Ile	agg Arg 240	gæt Asp	gcc Ala	JSS naK	aga Arg	aga Arg 245	Arg	cct Pro	gat Asp	gat Asp	ece Pro 250	891
ctt Leu	tac Tyr	gat gat	aga Arg	aag Lys 255	acc Thr	tta Leu	cac Bis	ata Ile	cca Pro 260	cct Pro	gat Asp	gtt Val	ttc Phe	aag Lye 265	aaa Lys	939
atg Met	tct Ser	Sca Sca	tca Ser 270	Caa Gln	ГÅЗ	çaa Çîn	tat Tyr	tgg Trp 275	agt Ser	gtt Val	aag Lys	agt Ser	gaa Glu 280	tat Tyr	atg Met	987
yab	att	gig Val 285	Leu	Phe	trt Phe	aaa Lys	gr.g Val 250	61 <i>à</i> 88 8	aaa Lys	ttt Phe	ta t Tyr	gag Glu 295	ctg Leu	tat Tyr	Glu gag	1035
cta Leu	gat Asp 300	g¢g Ala	gaa Clu	tta Leu	87À 8âţ	cac Mis 305	aag Lys	gag Glu	Leu	yab	12p 310	aag Lya	atg Met	acc Thr	atg Met	1003
5er 315	Gly	Val	GTÅ ââs	Lys	Сү <i>в</i> 320	Ary	Gla	Va1	Gly	11e 325	ser	Glu	ser	Gly	11e 330	1131
gat Asp	GJ n Gad	gca Ala	grg Val	caa Gln 335	PAB PSG	ctá Leu	tta Leu	gec Ala	cgt Arg 340	Gly gga	tat Tyr	aaa Lys	gtt Val	gga Gly 345	ega Axg	1179
atc Ile	Glu Glu	cag Gin	cta Leu 350	gaa Glu	aca Thr	tct Ser	yab 3gc	cat Gln 35 5	gca Ala	aaa Lys	Ala gec	yrg	ggt Gly 360	get Ala	aat Asn	1227
Thr	ata Ile	att Ile 365	610 ccs	agg	aag Lya	cta Leu	grt Val 370	cag Gln	gta Val	tta Leu	act Thr	cca Pro 375	tca 8er	aca Thr	gca Ala	1275
agr Ser	gag Glu 300	ejå ääs	āac Abņ	atc Ile	ggg Gly	cct Pro 385	gat Asp	gce Ala	gtc Val	cat His	ctt Leu 390	ctt Lau	get Ala	ata Ile	aa a Lyb	1323
gag Glu 395	atc Ile	aaa Lys	atg Met	gag Glu	cta Leu 400	caa Gln	aag Lys	tgt Cys	tca Şer	act Thr 405	gtg Val	tat Tyr	gga Gly	ttt Phe	gct Ala 410	1371
ttt Phe	gtt Val	gac Asp	tgc Cya	gct Ala 415	gcc	t t g Leu	agg Arg	ttt Phe	tgg Trp 420	gtt Val	gly ggg	tec Ser	atc Ile	agc Ser 425	gat Asp	1419

_	-				_			gcg Ala 435			_	_				1467
_	-				-			gjå 833				_	-			1515
_					_	_		937 832				_	_	_	-	1563
	_			-	_			aca Thr	-	_	_		_	-		1611
		_			_			aaa Lys	-			_				1659
_	•	-	_				-	car Cys								170 7
	•					_		agg				-	Азр			1755
_			_					caa Gln	-			Gly				1303
11e 555	Àsp	Gly	Gln	Thr	Met 560	Val	ນຮນ	Lev CCC	Glu	11e 565	Spe	АВЛ	Asn	Ser	Сув 570	1851
					Gly			tac Tyr		Tyr						1899
								agg Arg S95						Pro		1947
			Glu					a cgg					Ģlu			1995
		Asn					Ġζη	atc lle				IYI				2043

63!	5	, 201	, re/	, 610	640 Arg	Lau	. Leu	2 Gly	' Arg	f fle 645	Lye	Ser	s Sei	r Val	cga L Arg 650	2091
561	. 062	. 416	9 96I	655	Ç∻⊋ü	, tro	· vis	i Leu	Leu 660	GJA	Lys	Lys	val	669		2139
Glr	a oga a Arg	gto g Val	40a Lyg 670	W13	ttt Phe	GJA 333	caa Gln	att 1le 675	Val	Lys taaa	gly 399	tto Phe	aga Arg	Sei	gga Gly	2187
att Ile	gat Asp	ctg Leu 683	נו פיעו	ttg Leu	gct Ala	Cta Leu	69 0	Lys	gaa Glu	tca Ger	aat Asn	atg Met 695	Met	agt Ser	ttg Leu	2235
cet	tat Tyr 700	DAR	ctc Leu	tgt Cys	aaa Lys	ctt Leu 705	cct Pro	ata Ile	tta Leu	gta Val	gga Gly 71n	aaa Lys	age	gly	cta Leu	2283
gag Glu 715	Leu	ttt ?he	ctt Leu	tct Ser	caa Gln 720	t C C	gaa Glu	gca Ala	ηla	ata Ile 725	yab gsc	agc Ser	gac Asp	ttt	cca Pro 730	2331
aat Asn	tat Tyr	eja cag	23C A#N	caa Glp 735	даt	gtg Val	aca Thr	gat	gaa Glu 740	aac Aan	get Ala	Gln Gaa	act Thr	ctc Leu 745	Thr	2379
ata Ile	ctt Leu	ate Ile	gaa Glu 750	ctt Lev	ttt Phe	ate	gaa Glu	aga Arg 755	gca Gca	act Thr	caa Gìn	tgg Trp	tet Ser 760	gag Glu	gte Val	2427
att Ile	His	ace Thr 765	ata Ile	age Ser	tgc Cys	cta Leu	gat Asp 770	gtc Va).	ctg Leu	aga Ang	tct Ser	ttt Phe 775	gca Ala	ate Ile	gra Ala	2475
Ala	agt Ser 780	ctc Leu	tct Ser	gct Ala	gga Gly	agc 8er 785	Met	gcc Ala	agg	cct Pro	gct Val 790	att Sle	tct Phe	Pro ecc	gaa Glu	2523
tca Ser 79\$	Gja gaa	gct Ala	aca Thr	yat	cag Gln soo	aat Asn	cag Gin	aaa Lys	Thr	aaa Lyb B05	G1 y 333	cca Pyd	ata Ile	ctt Ceu	aaa Lys 810	2571
atc Ile	caa Gln	gga Gly	cta Leu	tgg Trp 815	çat His	eca Pro	ttt Phe	Ala	gtt Val 820	gca Ala	gc c	gat Asp	gly ggc	caa Gln B25	ttg Leu	2619
cct Pro	gtt Val	crg Pro	aat Asn 830	gat Asp	ata Ile	ctc Leu	Leu	99c 61y 635	gag Glu	gct Ala	aga Arg	Arg	agc Ser 840	agt Ser	gg¢ Cly	2667

WO 99/19492

											cca Pro					2715
		845		3			950			1		855	.,50	327	~~,	
						Ala					gtt Val					2763
	960					865					870					
											atc Ile					2811
875			- 4 -		380	•			-2.	885					890	
							-		_	_	atc	•				2859
Tnr	Ile	Phe	Thr	895	ren	Gly	Ala	Ser	900	Arg	Ile	Met	Thr	905	Glu	
_			_	_		_					tca	_		_		2907
Ser	Thr	Phe	310 Fea	Vai	Glu	Сув	Thr	915	Thr	Ala	Şer	Val	320 Een	Gln	Asn	
											crg					2955
λla	Thr	Gln 925	Asp	Ser	Leu	Va:	I∶e 930	Leu	Asp	61u	i.ev	Gly 935	yrg	Gly	Thr	
											gtt					3003
Şer	7hr 940	Phe	Ąsp	Gly	Tyr	945	Ilε	Ala	Туг	ser	Val 950	БрБ	Arg	His	Leu	
											acs					3051
955	GIQ	rys	vaı	GID	960	arg	Me t	Leu	Phe	965	The	8118	Tyr	нтв	970	
		_	_						•	•	ace Thr	_			-	3099
				975					980				•	985		
	-	-				_		_			CC it	_		-	_	3147
AT2	Cys	Ala.	990 990	rys	ser	¥td	zer	995 995	191	GIN	5 1.0		700 0	СУБ	Asp	
											gga					3195
GIN	_	Leu 1005	AST	bue	rea	_	arg 1016	ren	Thr	GIU	Gly	A1a 1015	-	Pro	Gia	
											ata					3243
	Tyr 1020	eth	Leu	Gln		Ala 1025	Leu	Met	Ala		Ile 1030	Pro	Asn	Olu	Val.	
gtt	988	aca	gca	tes	gat	øc.	act	caa	acc	ato	aag	202	tea	ath	703	3291
Val	Glu			Ser	Gly				Ala	Het	Lys			Ile	GLY	
103	5			;	104D					1045					1050	

			,	105	5	t ga r Gl	V LÆ	u Ar	g Se 106	r Gl O	u Ph	e Se	x Se	106	10 H. 5	ig	3339
			107	0	o 36	a tt	u va	107	у [] 5	8 Se	r Ar	g Va	1 Al 108	а Ні 0	s A	รก	3387
		108	5	e 01	, 91	n yai g da	109	р ту 0	r A9	o Th	r Le	u Ph 109	e Çy	s te	a to u Tr	.b 33	3435
141.	110	0	e ry	g Se	r se:	t tad r Tyr 1105	г Су: 5	9 Va∶	l Pr	o Ly	9						3478
																cta	3538
			atai	tatea	laa a	mari	١٢٢٠	נט ט:	cegai	ctaar	ં હાસ દ	23336	aaa	3333	aaaa	aaa	3598
aaa	laaaa	Ræ															3606
	1> .2> .3>		PR Ax	.09 T abid	lopsi	s th	erlis ve	na e	coty	pe c	'olum	bia					
< 40			31		PLIU	e no	πь .										
Met 1	Gln	. Arg	! Gln	Arg 5	ser	íle	Lev	. Ser	Phe 10		Gln	Lys	Pro	Thr 15		a	
Ala	Thr	Thr	Lys 20	gĵy	Гфл	Val	Ser	Gly 25	, YPD	Ala	Ala	Ser	Gly		613	•	
Gly	Ser	Gly 35	Gly	Pro	Arg	Phe	Asn 40	Val	Arg	Glu	Gly	Лар 45		Lya	Gly	,	
Asp	Ala 50	Ser	Va1	Arg	Phe	Ala 55	Val	ser	Lye	Ser	Val 60	Азр	Glu	Val	yrg	,	
Gly 65	Thr	дв£	Thr	Pro	Pro 70	Glu	Lys	Val	- Pro	Arg 75	Arg	Va1	Leu	Pro	9 0 1		
Gly	Phe	Lys	Pro	Ala 85	Glu	Ser	Ala	GJY	Авр 90	Ala	Ser	Ser	Leu	Phe 95	ser	•	
Aan	Ile	Met	eië OO <i>l</i>	Lys	Phe	Val	Lys	Val 105	Asp	Азр	Arg	A p	Cys 110	Ser	Gly	•	
Glu	Arg	ser 115	Arg	01u	Аѕр	Val	Val 120	Pro	Leu	asn	qeA	Ser 125	Ser	Leu	Суз		

Met	L75 130	Ala	Asn	Asp	Val	11e 135	Pro	Gln	Phe	yığ	Ser 140	Asn	Aan	e <i>ī</i> À	ГУЗ
Th¥ 145	61 n	Glu	Arg	Asn	Ні з 150	Ala	Phe	8er	Phe	Ser 155	Gly	Arg	Ala	Glu	ren Ten
Arg	Ser	val	Glu	Авр 165	Ile	Gly	Val	Asp	Gly 170	Двр	Val	PTO	Gly	Pro 175	Glu
Thr	Pro	Gly	Met 180	Arg	Pro	Arg	Ala	8er 185	Arg	Leu	ГÀЗ	Arg	Val 190	Leu	G) u
Авр	Glu	Met 195	The	Pµe	Lys	Glv	Asp 200	Lys	Val	Pro	Val	Leu 205	дар	Ser	Aan
Lys	Arg 210	Leu	Lys	Met	Ľеч	Gln 215	Asp	Pro	Val	Cys	Gly 220	Glu	Liys	ГÀВ	Glu
ขอ <u>1</u> 225	Asn	Glu	Gly	Thr	Ĺуs 230	Phe	Giu	дър	Len	Clu 235	Ser	Ser	уzg	Ιle	Arg 240
Asp	Ala	Asp	Arg	Arg 245	Arg	Pro	ASP)	Asp	250	L∌v	Туч	Азр	Arg	Lys 255	The
Lou	нэs	Ile	260	Pro	Asp	Val	Phe	L ув 265	Lys	Жet	Ser	Ala	Ser 270	Gln	Lye
Glp	Tyr	1'rp 275	Ser	Val	Lув	Ser	3В0 Сјл	Tyr	Met	Asp	Ile	Val 285	Leu	Phe	Phe
Lys	Val 290	Gly	Lys	Phe	Tyr	Glu 295	Leu	Tyr	Glu	Leu	Авр 300	Ala	Glu	ren	GJ?
Нів 305	Lув	Glu	Leu	Asp	Trp 310	Гув	Met	Thr	Met	Ser 315	_	Val	GJA	Lya	320
Arg	Gln	VAL	Gly	Ila 325		Glu	Ser	Glγ	11e 330	_	Glu	Ala	Val	Gln 335	Lys
Leu	Leu	Ala	Arg 340	GJλ	Tyr	Lys	۷al	Gly 345	_	Tle	Głu	Gln	leu 350	Glu	Thi
Ser	Asp	Gln 355	Ala	Lys	Ala	Arg	Gly 360		Aati	Thr	Ile	11 e 365		Arg	Ъуя
Leu	∀a1 370	Gln	Val	Leu	Thr	Pro 375		Thr	Ala	\$er	Glu 380		Asn	lle	Gly
Pro 385	Аєр	Ala	Val	His	Leu 390	Leu	Ala	lle	Lys	Glu 395		ГÀЗ	Met	Glu	Le:
Gln	Lys	СУВ	Ser	Thr 405		Tyr	Gly	Phe	Ala 410		Val	Авр	Суъ	Ala 415	

- Leu Arg Phe Trp val Gly Ser lie Ser Asp Asp Ala Ser Cys Ala Ala 420 425 430
- Leu Gly Ala Leu Leu Met Gln Val Ser Pro Lys Glu Val Leu Tyr Asp
 435 440 445
- Ser Lys Gly Leu Ser Arg Glu Ala Gln Lys Ala Leu Arg Lya Tyr Thr 450 455 460
- Leu Thr Gly Ser Thr Ala Val Oln Leu Ala Pro Val Pro Gln Val Met 465 470 475 480
- Gly Asp Thr Asp Ala Ala Gly Val Arg Asn Ile Ile Glu Ser Asn Gly 485 490 495
- Tyr Phe Lys Gly Ser Ser Glu Ser Trp Asn Cya Ala Val Asp Gly Lau
 500 505 510
- Asn Glu Cys Asp Val Ala Lou Ser Ala Leu Cly Glu Leu Ile Asn His 515 520 525
- Leu Ser Arg Leo Lys Leu Giu Asp Val Leu Lys His Gly Asp Ile Phe 530 535 540
- Pro Tyr Gln Val Tyr Arg Gly Cys Leu Arg Ile Asp Gly Gln Thr Met 545 550 555 560
- Val Aan Leu Glu Ile Phe Aan Aan Ser Cya Aap Gly Gly Pro Ser Gly 565 S70 S75
- The Leu Tyr Lys Tyr Leu Asp Asn Cys Val Ser Pro The Gly Lyc Arg
- Leu Leu Arg Asn Tro Ile Cys Bis Pro Leu Lys Asp Val Glu Ser Ile 595 600 605
- Asn Lys Arg Leu Asp Val Val Glu Glu Phe Thr Ala Asn Ser Glu Ser
- Met Gln Ile Thr Gly Gln Tyr Leu His Lys Leu Pro Asp Leu Glu Arg 625 630 635
- Leu Leu Gly Arg Ile Lys Ser Ger Val Arg Ser Ser Ala Ser Val Leu 645 650 655
- Pro Ala Leu Cly Lys Lys Val Leu Lys Gln Arg Val Lys Ala Phe
- Gly Gln Ile Val Lys Gly Phe Arg Ser Gly Il Asp Leu Leu Ala 675 680 685
- Leu Gln Lys Glu Ser Asn Met Met Ser Leu Leu Tyr Lys Leu Cys Lys 590 595 700

CONTINUE OF STREET

Сеч 705	Bip	lle	Гев	٧al	Gly 710	Lys	Ser	Gly	ГБЛ	Glu 715	ren	БрG	Leu	Şer	G1n 730
₽ħ e	G) u	Ala	Ala	Ile 725	Asp	Ser	Asp	Phe	Pro 130	Asn	Tyr	Gln	Asn	Gln 735	Λsp
Val	Thr	qeA	Glu 740	λеπ	Ala	Glu	Thr	leu 745	Thr	Ile	Leu	Ile	Glu 750	Leu	Phe
Ilė	Glu	Arg 755	Ala	Thr	Gln	Trp	Ser 760	Glu	Va1	lle	Kia	Thx 765	Ile	Ser	Cys
Гел	А sp 770	Vul	Leu	Arg	Ser	₽ħ₽ 775	A) a	Ile	Ala	Ala	ser 780	Leu	\$er	Ala	Gly
Ser 785	Met	Ala	Arg	Pro	Val 790	l)e	Phe	Pro	Glu	Ser 795	Glu	Ala	Thr	Asp	Gln 800
Aso	Gln	Lys	Thr	I.ys 805	αĵλ	Pro	1 e	Leu	Lys 810	Lle	Gln	Gly	Lev	Trp 815	Nis
Pro	शाल	Ala	Va). B20	Ala	λla	ysb	21 À	925 61n	Leu	Pro	٧æ١	Pro	780 088	Авр	Ile
Leu	Leu	Gly 935	Сјп	Ala	Arg	yrg	Ser 640	Ser	Gly	Ser	Ile	His 645	Pro	Arg	Ser
ren	650	Leu	Thr	Gly	Pro	Asn 855	Met	Gly	Cly	Lys	S⊕r 860	Thr	Len	Теп	Arg
Ala 865	Thr	Cys	Leu	Ala	Val 870	Ile	Phe	Ala	Gln	675	gly	СУя	туг	Val	880 Bro
Сув	Glu	Ser	Cys	Olu BBS		ser	Leu	Val	ASP 0ea	Thr	Ile	Phe	Thr	Arg 395	Leu
Gly	Ala	Ser.	Asp 900	Arg	Ile	Met	Thr	Gly 905	Glu	Ser	Thr	Phe	Ն ը 910	Val	Glu
Сув	Thr	Glu 915	Thr	Ala	Ser	Val	Leu 920	Gln	Asn	Ala	Thr	Gln 925	Asp	Ser	Lev
Val	11e 930	Leu	Asp	Glu	Leu	Gly 935	Аrg	Gly	Thr	Ser	Thr 940		Ąsp	Gly	Tyr
Ala 945	Ile	Ala	Тут	Ser	Val 950	Phe	Arg	His	Leu	Val 955	Glu	ГЛЯ	Val	Gln	Cys 960
Arg	Met	Leu	Phe	Ala 965	Thr	Жiв	Tyr	BŹĦ	Pro 970	Leu	Thr	Lys	Glu	Phe 975	Ala
ser	Ķiş	Pro	Arg	Val	Thr	Ser	Lys	His	Met	Ala	ÇAs	Ala	Рие	Lys	Ser

Arg Ser Asp	Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu	
	1000 1005	
1010	Thr Glu Gly Ala Cys Pro Glu Sen Tyr Gly Leu Gln Val 1015 1020	
Ala Leu Met 1025	Ala Gly Ile Pro Asn Gln Val Val Glu Thr Ala Ser Gly 1030 1035 1040	
Ala Ala Gln	Ala Met Lys Arg Ser Ile Gly Glu Asn Phe Lys Ser Ser 1045 1050 1055	
Glu Leu Arg	Ser Glo Phe Ser Sør Leu His Glu Asp Trp Leu Lys Sør 1060 1065 1070	
Leu Val Gly 1075	Ile Ser Arg Val Ala Mis Asn Amo Ala Pro Ile Gly Glu 1080 1085	
Asp Asp Tyr 1090	Asp Thr Lau Phe Cys Leu Trp His Glu IIc Lys Ser Ser 1095 1100	•
Tyr Cys Val	Pro Lys	
. 77.0		
<210>	32	
<211>	24	
<212>	ONA	
<213>	Artificial sequence	
<220>		
<233>	Forward primer for PCR amplification of ATRGENEA microsatellite	
<400>	32	
accatgoata g	CTtaaactt cttg	24
<210>	33	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<320>		
<223>	Reverse primer for PCR amplification of ATHGENEA microsatellite	
<400>	33	
acataaccac a	Bataggggt oc	22

בודה ארם בסום אסא אידו

. \$.

<210>	34			•
<211>	18			
<212>	ona			
<213×	Artificial sequence			
<220>				
<223>	Porward primer DMCIN-A for PCR	on genemic	DNA o	f Arabidocsis
	thaliana ssp. Landsberg eracta	"Ler"		
<400>	34			
gaagogatat	cqttcgtg			18
<210>	35			
<211>	19			
<212>	DNA			
<213>	Artificial sequence			
(213)	Artificial sequence			
<220>				
<223>	Reverse primer DMCIN-B for PCR thaliana ssp. Landsberg erecta		DNA o	f Arabidopaja
<400>	35			·
agar tọcg aợ	aacattcc			13
			·	
<210>	36			
<211>	31			
<212>	DNA			
<213>	Artificial sequence			
<220>				
<223>	Forward primer DMCIN-1 for PCR thaliana map. Landaberg erecta		DNA o	f Arabidopsis
<400>	36			
acgcgtcgac	tcagetatga gattactcgt g	•		31.
<210>	37			
<211>	29			
<212>	DNA			
<213>	Artificial sequence			
<220>				
<273>	Reverse primer DMCIN-2 for PCR	on denomic	DNA o	f Arabidopsis
	thaliana ssp. Landsberg erenta	_		- 1.2 a - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2
<400>	37			
getetagate	teteqeteta agactetet			29

```
<210>
               38
  <211>
               32
  <212>
               DNA
  <2135
               Artificial sequence
 <2205
  د2233ء
              Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsia
              thaliana ssp. Landsberg erecta "Ler"
 <400×
              38
 getetagage tretettaag taagtgattg at
                                                                            32
 <210>
              39
 <211>
              48
 <212%
              DNA
 <213>
              Artificial seguence
 <2205
 <223×
              Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis
              chaliana sap. Landaberg erecta "Ler"
 <400>
              39
 tootoogggs togagagate tecatggitt circagerer atgastee
                                                                           48
<210>
              40
<2115
              26
<2125
              DNA
<2135
              Artificial sequence
<220>
<223>
             Forward primer DMC1a for PCR on genomic DNA of Arabidopeis
             thaliana asp. Landsberg erects "Ler"
<400>
             40
acgegtegae gaattegeaa gtgggg
                                                                           26
<210>
             41
<211>
             38
<212>
             DNA
<213>
             Artificial sequence
<220>
<223×
             Reverse primer DMClb for PCR on genomic DNA of Arabidopsis
             thalians sap. Landsberg erecta "Ler"
<400≈
             41
```

fccatggaga	totocogggt acceatotge trogaggg	38
<210>	42	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>	-	
<223>	Forward primer for PCR amplification of ATEAT1 SSLP marker Arabidopsis thaliana subapecies	ir
<400>	42	
gccactgcgt	gaatgatatg	20
-210.	•	
<210×	4.3	
	22	
-212>	DNA	•
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of ATEAT1 SSLP marker	- 1 m
	Arabidopsis thaliana subspecies	
<400>	43	
ogaacageca	acattaattc cc	22
<210>	44	
<311>	18	
<212>	DNA	
<313>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA63 SSLP marker	in
	Arabidopsis thaliana subspecies	• • •
<400>	44	
aaccaaggca	cagaagcg	18
	·	
<210>	45	
<211>	18	
<212>	DNA	
<213>	Artificial sequence	
<220×		
<223>	Reverse primer for PCR amplification of NGA63 SSLP marker	in
	Arshidonais thallans subspecies	~

```
<400>
                45
   acceaagiga tegerace
                                                                              18
  د210ء
                46
  <211>
                21
  <212>
                DNA
  <213>
               Artificial sequence
  <220>
  <233>
               Forward primer for PCR amplification of NGA248 SSLP marker in
               Arabidopsis thaliana subspecies
  <400>
               46
  tассувасся азасасваа<u>д</u> д
                                                                             21
 <210>
               47
 <211×
               22
 <212>
               DNA
 <2135
               Artificial sequence
 <220>
 <223>
              Reverse primer for PCR amplification of NGA248 SSLP marker in
              Arabidopsis thaliana subspecies
 <400>
               47
 totgtatoto ggtgaattot co
                                                                            22
 <210>
              48
 <2115
              22
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
              Porward primer for PCR amplification of NGA128 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              48
ggtctgttga tgtcgtaagt cg
                                                                           22
<210>
              49
<211>
             22
<212>
             DNA
<213>
             Artificial sequence
<220>
```

WO 99/19492 PCT/EP98/06977

<2235	Reverse primer for PCR amplification of NGA128 SSL Arabidopsis thalians subspecies	P marker i
<400>	49	
atottgaaac	r ctttagggag gg	22
<210>	50	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223×	Forward primer for PCR amplification of NGA280 SSL Arabidopsis thaliana subspecies	P marker in
<400>	So	
ctgateteac	c ggacaavagr gc	22
	·	•
<310>	51	
<211>	20	
<212>	DNA	•
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA280 SSL Arabidopsis thaliana subspecies	P marker in
<400>	51	
ggctccataa	a sasdcâtecc	20
<310>	5 2 .	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>	•	
<223>	Forward primer for PCR amplification of MGAlll SSL Arabidopsis thaliana subspecies	P marker in
<400>	52	
ctccagttgg	j aagstaaagg g	31
<210>	53	
<211>	21	
<212>	DRA	
4213P	VNA	

<220>	
<223>	Reverse primer for PCR amplification of MGA111 SSLP marker in Arabidopsis thaliana subspecies
<400>	53
tgtttttag	g gacaaatgge g
	21
<210>	54
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<333>	Forward primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	54
ccttcacato	- Caaaanneesc 20
<210>	\ \$5
<3115	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGALEB SSLP marker in Arabidopsis thaliana subspecies
<4005	55
gcatataccc	acaaccagaa 20
<210>	56
<211>	. 20
<212>	DNA
<213>	Artificíal sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis thatiana subspecies
<400>	56
cgctacgctt	ttoggtasag 20

PCT/EP98/06977

WO 99/19492 PC

<210>	57
<211>	20
<212>	DNA
c2135	Artificial sequence
<220>	
€2235	Reverse primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis chaliana subspecies
<400>	57
goacagteca	agteacaace 20
<210>	58
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	•
<223≽	Porward primer for PCR amplification of NGA361 SSLP marker in Arabidopsis Chuliana subspecies
<400>	5.6
aaagagatga	gaatttggar 20
<310≻	59
<211>	23
«212»·	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA361 SSLP marker in
~~~	Arabidopsis thaliana subspecies
<400>	59
acatatcaar	atattaeagt agc 23
<210>	6.0
<211>	18
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA168 SSLP marker is Arabidopsis thaliana subspect s
<400>	60

ceattracts	Cactingo
	18
<210>	61
<211>	32
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Develop Things for the
	Reverse primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	61
gaggacatgt	ataggageet eg 22
	22
<210>	62
<211>	20
<212>	DNA
<213>	Artificial sequence
	and And the state of the state
<220>	
<223>	Porward primer for PCR amplification of Athbroz SSL2 marker
	in Arabidopsis chaliana subspecies
<400>	62
tgacctcctc t	Pagagone v
-2	20
	•
<210>	63
<211>	22
<212>	DNA
(Z13>	Artificial sequence
:320>	
:223>	Pavavos primar for an
	Reverse primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thatians subspecies
400>	63
taacagaaa c	ccaaagett te
210>	64
211>	21
212>	DNA
213>	Artificial sequence
220>	
223>	Porward primer for PCR amplification of AthURIQUE SSLP marker
	in Arabidopsis thaliana subspecies

4400s	64	
aggcaaatgt	coattteatt g 21	L
<210>	65	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<2233	Reverse primer for PCR amplification of AthUBIQUE SSLP marke in Arabidopsis thaliana subspecies	ıΥ
<400>	65	
acgacatggc	agatttetee 20	>
<310>	66	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA172 SSLP marker i Arabidopsis thaliana subspecies	ת.
<400>	66	
agetgettee	ttatagegte c 21	•
<210>	67	
<211>	19	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA172 SSUP marker i	13
	Arabidopsia thaliana subspecies	**
<400>	67	
catoryaatg	ccattgt:c	i
<210>	68	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<230>	$\cdot$ .	

```
< 223×
                 Forward primer for PCR amplification of NGA136 SSLP marker in
                 Arabidopsis chaliana subapecies
   <400×
                 58
   gasaaaacgo tactttcgtg g
                                                                              21
   <210>
                59
   <211>
                22
   <212>
                DNA
   <2113>
                Artificial sequence
   <220>
   <223>
                Reverse primer for PCR amplification of NGA126 SSLP marker in
                Anabidopsis thalisma subspecies
   <400>
                69
  caagagcaat atcaagagca gc
                                                                             32
  <210×
                70
  <211>
                23
  <212>
                DNA
  <213>
                Arcificial sequence
  <220×
               Forward primer for PCR amplification of NGA162 SSLP marker in
  <223>
               Arabidopsis thaliana subspecies
  <400>
               70
  catgeaattt geatergagg
                                                                            20
  <210>
               71
  <211>
               22
  <212>
               DNA
  <213>
               Artificial sequence
 <220>
 <223>
               Reverse primer for PCR amplification of NGA162 SSLP marker in
               Arabidopsis thaliana subspecies
 <400>
               71
 ctctgtcact cttttcctct gg
                                                                            22
 <210>
               72
 <211>
               21
 <212>
              DNA
<213>
              Artificial sequence
```

WO 99/19492 PCT/EP98/06977

<220>		
<3235	forward primer for PCR amplification of NGA6 SSLP marker Arabidopsis thaliana subspecies	in
<400>	72	
tggatttctt	cctctcttca c	21
<210>	73	
<211>	21	
<212>	DNA.	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA6 SSLP marker Arabidopsis thaliana subspecies	in
<400>	73	
atggagaagc	thacactgat c	31
s210>	74	
<211>	26	
<212>	DNA	
<213>	Artificial sequence	
42172	Altilicial seguence	
<22B>		
<2235	Forward primer for PCR amplification of NGA12 SSLP marker	io
	Arabidopsis thaliana subspecies	111
<400>	74	
aatgttgtcc	tececteete	20
	•	
<310>	75	
<211>	22	
<312>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA12 SSLP marker Arabidopsis thalians subspecies	in
<400>	75	
tgatgctctc	CQBBSCBBQB QC	22

<210>	76
<211>	21
<212>	DNA
<213>	Artificial sequence
	and reguence
<520>	
<223>	Forward primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	76
gagggcasat	cttratrog g
<210>	
<211>	77
<212>	22
<213>	DNA
~2137	Artificial sequence
<220>	
<323>	Paulane must
	Reverse primer for PCR amplification of NGAB SSLP marker in Arabidopsis chalisms subspecies
<400>	77
tggetttegt	ttatasacat cc 22
<210>	78
<211>	21
<212>	DNA
<313>	Artificial sequence
:2205	·
223>	Power and and a second
	Forward primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
400>	78
cgaaaaaac .	aaaaaatee a
310>	79
211>	21
212>	DNA
213>	Artificial sequence
220>	
223>	Reverse primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thalians subspecies
100>	79

cdacâ#sccd	ACAGAALEAG G	21
<210>	8 C	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	•
<320>	•	
<223>	Forward primer for PCR amplification of NGA225 Arabidopsis thaliana subspecies	SSLP marker in
<400>	30	
gaeatccaaa	tcccagagag g	21
<210>	Bi	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	•
<220>		
<223>	Reverse primer for PCR amplification of NGA225 Arabidopsis thaliana subspecies	SSLP marker in
<400>	91	
totocccact	äğttttgtgt ce	<b>. 2 2</b>
<210>	B2	
<311>	19	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA249 and Arabidopsis thaliana subspecies	ssip marker in
<400>	82	
taccgtcaat (	ttcatcgcc	19
<210>		
<211>	83 22	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<2 <b>2</b> 3>	Davance seimer for son	
~ <b>~ ~</b> ~ ~	Reverse primer for PCR amplification of NGA249 and Arabidopsis thaliana subspecies	SSLP marker in

```
<400>
                  83
    ggatcoctaa ctgtasaatc cc
                                                                               22
   <210>
                 84
   <211>
                 22
   <212≥
                 DNA
   <213>
                 Artificial sequence
   <220>
   <223>
                Porward primer for PCR amplification of CA72 SSLP marker in
                Arabidopsis thaliana subspecies
   <4005
                84
  satocoagta accasacaca ca
                                                                              22
  <210s
                85
  <3115
                20
  <212>
                DNA
  د215×
               Armificial Requosce
  <220>
  <223>
               Reverse primer for PCR amplification of CA72 SSLP marker in
               Arabidopsis challana subspecies
 <400>
               85
 cceagtetaa ccacgaccac
                                                                            20
 <210>
               86
 <2115
              20
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
 <2235
              Forward primer for PCR amplification of NGA151 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              86
gtttcgggaa gttttgctgg
                                                                           20
<210>
             87
<211>
             24
<212>
             DNA
<213>
             Artificial sequence
<220s
```

<223»	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA151	sslp	marker	in
<400>	8.3		•			
cagtotadas	gogågågtat gatg				•	24
<310>	B8					
<211>	22					
<212>	DNA					
<213>	Artificial sequence					
<220>	·					
<223 >	Forward primer for PCR amplification	~ <del>•</del>	NORTH	601.6		
	Arabidopsis thaliana subspecies	O1	NGATOS	SOLP	MAZMET	111
<4DQ>	88					
gttatggagt	ttctagggea eg				•	22
						•
<210×	89					
<211>	20					
<212>	ANG					
<213>	Artificial sequence					
<230>						
<223 ×	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA106	SSLP	marker	in
<400×	B9					
tgccccateu	tgttcttctc					20
<210>	90	_				
<211>	20					
<212>	DNA					
<313>	Artificial sequence					
<220>						
<223>	Forward primer for PCR amplification Arabidopsis thalians subspecies	of	NGA139	SSLP	marker	in
<400×	90					
agagccacca	gatccgatgg					20
.750			4		1	
<210>	91					
<211>	21					
<212>	DNA					
<213>	Artificial sequence					

<220»		
<223>	Reverse primer for PCR amplification of NGA139 SSLP marker Arabidopsis thaliana subspecies	in
<400>	91	
ggtttegtt	tt cactatocag g	21
<210>	93	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	4
<220>		
<223>	Forward primer for PCR amplification of NGA76 SSLP marker Arabidopsis thaliana subspecies	iπ
<400>	92	
ggagaaat	g tcaccctcca cc	22
<210>	93	
<21.1>	20	
<212>	DNA	
<213>	Artificial sequence	
<230>		
<223>	Reverse primer for PCR amplification of NGA76 SSLP marker t Arabidopsis thaliana subspecies	in
<400»	93	
aggcatggga	a gacatttacg	20
<210>	94	
<211>	. 20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATHSO191 85LP marke in Arabidopsis thalisma subspecies	r
<400>	94	
ctccaccaat	catgessatg *	0

бD

	A.B.
<210>	95
<211>	21
c212>	DNA
<213>	Artificial sequence
c220>	
<223>	Reverse primer for PCR amplification of ATHSO191 SSLP marker
	in Arabidopsis thaliana subspecies
<400>	95
tgatqttgat	ggagatggtc a 21
<210>	96
c211>	22
<212>	DNA
<213>	Artificial sequence
<220>	•
<223>	Forward primer for PCR amplification of NGA129 SSLP marker is
	Arabidopsis thaliana subspecies
<400>	96
tcaggaggaa	ctaaagtgag gg
<210>	97
<311>	22
<212>	DNA
<213>	Artificial sequence
45207	wifter reducite
<230>	
<223>	Reverse primer for PCR amplification of NGA129 SSLP marker in Arabidopsis thaliana subspecies
-450.	
<400>	97
cacactgaag [.]	atggtcttga gg
<210>	
<211>	98
<211>	\$062
	DNA
<213>	Arabidopsis thaliana ecotype Columbia
<220>	
<223>	Ganomic DNA sequence of AtMSR6
<400>	97
ztttttggtt	get acaata asggtatace gttttatete atem tatam ctatatataa 60

aagaaasgaa ayabababab tgetttttcca				120
ttacttttta çattqqtcaa casaatacaa	a gatasacga	c alognotas	it catttessaa	130
Ctttaccott aagtttaada tetagaadot	tetecatet	t egcaageac	a geetgattag	240
gaacagettt accattetea tatteetgaa	t ctacctyag	t cototratt	g atctgtttcg	300
ccaaatcegc ttgtgacatc ttcttctcca	atctogoto	t cognaticat	¢ aåbotcacet	360
chgetthead adgatecate googeagget	ctgtttctt	ttccagett	c ttogtgttaa	420
tdadoggaad ogcogtagat ttoccottbb	tgttcgaac	ggcalogaa	t ttottaaccg	480
tttgaaccgc gacaccgttt ctcagagctg	cgttaacege	tteeggate	g cgtaggtett	540
ggctettttg ttttgatttg tggagaacta	ctggttccc	gtettgtgt:	t actgctcetg	<b>600</b>
ggtatetget eggeategte gatgaattga	gagaaaggaa	caacgcgaa:	a attttattaa	ಕ್ಕಳು
tetgageett gaaattgaga aacgatgaag	acgaagaatg	tigttgaga:	gattqtgata	720
tttatacata cgaagattgg tttctggaga	attegateac	ರಚಿತ್ರವಾಗಿದ್ದ	attttegtet	780
ctggaacgte cttagagatg attgacgacg				840
ctttttgggt tggattegtg gtacaccata				900
aattoggott toggttoggt tatgagtoat				960
taaaatttat tteggtttea atteggatee	cttacttcca	gaaccgaatt	attegaaace	1020
ggggttagee gaaccgaata ccaatgceeg	attgastegs	tggotagaaa	gatecaacgg	1080
tatacsatsa tagancatas atoggacgg,	catexaagee	tcaaagagtg	sacagtcasc	1140
aaaaaaagtt gagccctgag gagtatcgtt		•		1200
ttttggcgcc aatotttooc coctttogaa				1260
cectototot cacaattoca aaaaatgcag	cgccagagat	cgattttgtc	tttcctccaa	1320
asacccaegg eggegaetae gaagggtttg	gtttccggcg	atgetgetag	çããçãããããc	1380
ggcagcggag accacgattt aatgtgaagg	aaggggatgc	caaaggcgac	gettetgtae	1440
sttttgctgt ttcgaaatct gtcgatgagg	ttagaggaac	ggatactoca	ccdardardd	1500
steegegteg tgteetgeeg tetggattta .	ageeggetga	atcegceggt	gatgettegt	1560
cotgitete caatattaty cataagititg	teaaagtcga	tgatogagat	tgttctggag	1620
gaggtacta atettogatt ctcttaattt (	tgttatettt	agetggaaga	agaagatteg	7683

tgtaatttgt	cgtattcgtt	ggagagatte	tgattactgc	attggatogt	tgtttacaaa	1740
titttcaggag	cogagaagat	gttgttccgc	tgaatgatec	atctctatgt	atgaaggota	1800
atgatgttat	teetraattt	cgttccaaca	atggtaaaac	tcaagaaaga	aaccatgott	1860
ttagttteag	tgggagaget	gaacttagat	cagtagaaga	tataggagta	gatggcgatg	1920
cteatggtes	agaaacacca	gggatgegte	cacgtgcttc	tegettgaag	cgagttctgg	1980
aggatgaaat	gacttttaag	gaggataagg	tteetgtatt	ggactctaac	aaaaggctga	2040
aaatgctcca	ggatccggtt	tgtggagaga	agaaagaagt	aaacgaagga	accasatttg	2100
aatggettga	gtettetega	atcagggatg	ccsatagaag	Acgtoctgat	gatcecettt	2160
acgatagaaa	gaccttacac	ataccacctg	atgttttcaa	gasaatgtct	gcatcacasa	2220
agcaatatcg	gagtgttaag	agtgaatata	tggacattgt	gettttettt	aaagtggtta	2280
gtaactatta	atctagtgtt	caatccatct	cctcaatgcg	atttgttcac	ttacatotgt	2340
tracgitarg	CtCttCtCag	gggaaatttt	atgagctgta	tgagctagat	gcggaattag	2400
gtcacaagga	gcttgactgg	aagacgacca	tgagtggtgt	gggaaaatgc	agacaggtaa	2460
attagttgaa	acaactggcc	tgottgaatt	attgtgtcta	taaattttga	caccaccttt	2520
tgtttcaggt.	tggtatetet	gaaag¢ggga	tagatgaggc	agtgcaaaag	ctattagete	2580
ardaaraada	gaaccatcat	actttatgga	attogtttac	tgctacttcg	gctaggattt	2640
aagaaatgga	aatcacttca	agcatcatta	gttaggatcc	tgagaactca	ggatgttttc	3700
ttättegtta	tataataagt	cttttcatca	aggagtaaca	aacaaaactt	gcacaatatt	2760
tgtgtgctca	ctggcaaggr	atatatacce	agctaacctt	tgctagttca	ctgtagtaac	2820
agttacggat	aatatatgit	tacttgtatg	tggtaccetc	attitgtete	tcatggaggc	2880
tttcaagcct	tgigttgasa	ctggacagtt	acatațgott	ccaacagaaa	ctagcatgca	2940
gattcatatg	ctttcctatt	ctactaatta	tgtattgaca	cactcgttgt	ttcttttgaa	3000
agatataaag	ttggacgaat	cgagcagcta	gaaacatctg	accaagcaaa	agccagaggt	3060
gctaatactg	taagttttct	tggataggtc	aaggagagtg	ttgcagactg	tttttgatca	3120
tttcttttc	tgtacattac	tttcatgctg	taattaactc	aatggctatt	ctggtctgat	3180
tatcagataa	ctccaaggaa	gctagttcag	gtattascte	catcaacage	aagegaggga	3240
aacatogggo	ctgatgecqt	CCRTCTTCTT	gctataaaaq	Aggtttgttz	tttactratt	3300

tatettatea tetteagtte atecangtes but	
tatettatea tetteagtte atecasete teaasastta esetettett taecastett	3360
ceatcaaget gegtaaagga titggaatta gaaaateatt attigatget tegtettata	3420
vycaagaggt tecettgaaa agatetgtt aagattettt geaettgaaa aatteaatet	3480
tettaagtga atcreetact teettacaat gateatagte tgesattgea tgtesagtsa	
tateatteet tyttactyca tececetett tettaatyae cattytetat yttytytty	3540
totogtgtgd tggagaaaat gatagotgat ocaagotgta cattatoatg attaagtago	3600
tgotcaggaa ttgoctttgg ttacatroo tantamin	3660
tgercaggaa ttgeetttgg ttacattgee taatggtttg atgteaattt ttettetgaa	3720
tetttattt agateaaaat ggagetacaa aagtgtteaa etgtgtatgg atttgetitt	3780
gttgaetgtg etgeettgag gttttgggtt gggtccatca gegatgatge atcatgtget	3840
getetiggag egitatigat geaggiaage aagiglatte iglatettat giglaceatg	3900
egactecetg tgestatatt tgggetgeag gaactaatte tgaateacea tetggtatgt	3960
tttttreagg tttrteraaa ggaagtgtta tatgacagra aaggtaaact gettgtateg	4020
ccagtegtet tgttaaacag aatstaaggt aaatgacact ggttaattta aagtgcatac	
atgitgaaat atigoagggo tatoaagaga agoacaaaag gototaagga aatatacgot	4080
gacaggtace attteagtag graagetaae tgacaattta accycteace gaatgatagg	4140
retertabae attgetaatg tagatgatgt ttatgtttea atetaatagg gtetaeggeg	4200
gtacagetgg etceagtace against and a	4250
gtacagtigg ctccagtacc acaagtaatg ggggatacag atgctgctgg agttagaast	4320
ataatayaat ctaacggata ctttaaaggt tcttctgaat catggaactg tgctgttgat	4380
ggtctaaatg aatgtgatgt tgreettagt getettggag agetaattaa teatetgtet	4440
aggetakagg tgtgttgget tgtttagttt ttgettttca caaattaage aaaggaactt	4500
ttestaaett aeagttteta tetaettgea getagaagat gtaettaage atggggatat	4560
ttttccatac caagtttaca ggggttgtct cagaattgat ggccagacga tggtasatct	
tgagatattt aacaataget gtgatggtgg teetteagge aagegeatat ttetettttg	4620
ataacttcaa ctagagggca gacatagaag gaaaaattct.aatacttcgt acggatctcc	4680
ageasgraat ageogatett torreages has	4740
agtazgtaat agcogatttt tgtttaccta tgtagggacc ttgtacaaat atottgataa	4800
ctgtyttagt ccaactggta agegactett aaggaattgg atetgecate caetcamaga	4860
tgtagaaago atcaataaac ggcttgatgt agttgaagaa ttcacggcaa actcagaaag	4920

WO 99/19492

tatgcaaate	actggccagt	atctccacaa	acttccagac	ttagaaagac	tgctdggacg	4950
catcaagtet	agcgttcgat	catcagecte	tgtgttgcct	getettetgg	ggaaaaaage	5040
gotgaaacaa	cgagtaagta	tcastcacas	gttttetgag	taatgrette	catgagtagt	5100
acaggactaa	aacattacgg	gtotagetaa	agactgttct	cottottttg	caatgtctgg	5160
ttattcatta	catttetett	aacttattgc	attgcaggst	aaagcatttg	ggcasattgt	5220
gaaagggttc	agaagtggaa	ttgatctgtt	gtrggctcta	cagaaggaat	Czaatatgat	5280
gagtttgctt	tatasactct	gtsaacttcc	tatattagta	ggasasaggg	ggctagagtt	5340
attotttot	caattcgaag	cagceataga	tagcgacttt	ccaaattat¢	aggtgcccat	5400
ctatctttca	tacettacaa	caaaatgtct	gtcactactc	aaagcaatge	atatgg¢tta	5460
gateteaaet	cacaccccga	ggatcctaaa	gggatttg¢t	tittatteet	aatgtttttg	5520
gatggtttga	tttatttcta	acttgaactt	attaatettg	taccagaacc	aagatgtgac	5580
a <b>gatgaaaa</b> c	gccgaaactc	tcacaatact	tatcgaactc	tttatogaaa	gagcaactca "	5640
atggtotgag	gtcattcaca	ccataagctg	cctagatgtc	ctgagatctt	ttgcaatcgc ,	5700
agcsagtctc	tctgctggaa	geatggecag	gcctgttatt	tttcccgaat	cagaagetac	5760
agatcagaac	cagaaaacaa	aagggccaat	acttaaaatc	caaggactat	ggcatccatt	5820
tgcagttgca	gccgatggtc	aattgeetgt	tccgaatgat	atactecttg	gegaggetag	SBBO
aagaagcägt	ggcagcattc	atcctcggtc	attgttactg	асдадассаа	acatqgqcqg	5940
aaaatcaact	cttcttcgtg	caacatgtet	ggeegttate	tttgcccaag	tttgtatact	6000
cgttagataa	ttactctatt	ctttgcaatc	agttcttcaa	catgaataat	aaattetgtt	6060
tratgtatga	agcttggctg	ctacqtgccg	tgtgagtett	gcgaaatctc	cctcgtggat	6120
actatcttca	caaggcttgg	cgcatctgat	agaat catga	caggagagag	taagttttgt	6180
tctcaaaata	ccaattcctc	gaactattta	ctcagatttt	gtetgattgg	acaaggtggt	6240
trrgortttt	tttaggtacc	tttttggtag	aatgcactga	gacagcgtca	gttcttcaga	6300
atgcaactca	ggattcacta	gtaatccttg	acgaactggg	cagaggaact	agtactttcg	<b>636</b> 0
atggatacgc	cattgcatac	toggtaacet	getettetee	ttcaacttat	acttgttgat	6430
Caacraaaac	atgcaattca	ttttgctgaa	acttattgat	ttatatcagg	tttttcgtca	6490
cctggtagag	aaagttcaat	greggatget	ctttgcaaca	cattaccacc	ctctcaccaa	6540

ggaattogog totoaccoac gigtoaccio gaaacacatg gotigogcat teaaatcaag	6600
atotgattat caaccacgig grigigatca agacctagig ticitgiacc giblaaccga	
gggagettgt eetgagaget aeggaettea agtggeacte aeggetggaa taccaaacea	6 <b>5</b> 60
agregatigas acagestes graciactes agreetas agatematic gegansacti	6720
caagtcaagt gagctaagat ctgagttoto aagtotgcat gaagactggo tcaagtcatt	67B0
ggrgggtatt totogagtog occasages traces.	€840
ggrgggtatt totogagtog orcacaacaa tgodcocart ggcgaagatg actacgacac	6900
ttlgttttgc ttatggcatg agatcaaatc ctcttactgt gttcccaaat aaatggctat	5960
gacaraacac tatorgaago togitaagio tittgottot otgatgitta ticotottaa	7020
aaaatgotta tatatoaaaa aactgtttoo togattataa caagattata tatgtatotg	7080
toggtttage tatggtatat aatatatgta tgttcatgag attggtcaag agaaatacto	7140
acaaacagta tattaagaag gaaatatgtt tatgcattaa tttaagtttc aagataaact	7200
gcazataacc tegactadag ttgcaaagae caaacacaaa ttacaaaact tataagaett	7260
aagitetgaa tteectaaaa eeaaaaaa aaacagaaca tattttgttg eatetacaaa	7320
caacacaaac ctacatagtt tataacttac tcatcactga gattaacatc agaaccattc	7380
tecatetett catetteact eteateatea teaceaceae catgatgatt etectectet	7440
tracgtaace tagcastote actotgaget ctatesaces tetgettett etgesactee	7500
adatototet gaadatrage totoatette tecaseteet teattegete tetettacte	7560
ttetecatet teteataaac etteccaaac eteteaacag aateegecaa catettatae	
gaageagegt cattaacett ettecteteg tacteaacet catcatecte atectectee	7620
tolteagaat caccaggact atocatcate teatcadace cattagactt atotaaataa	7680
Accitagigt teatagacae saacteasst manage cattagacit atetagatas	7740
accitagigt teatasacae saacteacet gaateaacae cacaagetaa acciasatee	7800
gacttyggcg aaacacaaag caacatatec aacttattga aaaacgacca tttacttgaa	7860
dotaaacotg atticicaac citaatetic tottibolat Acticolott caagicatea	7920
abdattotoc tacattgogt otragattto tocatoetta gotocoodoc cactttotra	7980
gotacticat tecassecte gsteetcaaa etecttotae eeaattgeaa aaacetatet	8040
CCCAAACII daagcaacad aa	0062

### **PCT**

#### WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



### INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/82, 15/29, C07K 14/415, C12N 15/10, 5/04, 5/14, A01H 5/00

**A3** 

(11) International Publication Number:

WO 99/19492

(43) International Publicati n Date:

22 April 1999 (22.04.99)

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

ΑU

(71) Applicant (for all designated States except US): RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DOUTRIAUX. Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40 Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux, F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).

(74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

4

(88) Date of publication of the international search report. 24 June 1999 (24.06.99)

### (54) Title: METHODS FOR OBTAINING PLANT VARIETIES

### (57) Abstract

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

# FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL AM AT AU AZ BA BB BE BF BG BJ BR CCF CG CH CI CM CU CZ DE DK EE	Albania Armenia Austria Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Côte d'Ivoire Cameroon Chima Cuba Czech Republic Germany Denmark Estonia	ES FI FR GA GB GE GH GN IE IIL IS IT JP KE KG KP KR LC LI LK LR	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia	LS LT LU LV MC MD MG MK ML MN MR MW MX NE NL NO NZ PL PT RO RU SD SE SG	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden Singapore	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovakia Scnegal Swaziland Chad Togo Tajikistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam' Yugoslavia Zimbabwe
--------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------

Inte. onal Application No PCT/FP 98/06977

C 4 61 4 6 6			CI/EP 98/06977
IPC 6	HFICATION OF SUBJECT MATTER C12N15/82 C12N15/29 C07K14/ C12N5/14 A01H5/00	/415 C12N15/10	C12N5/04
According (	to International Patent Classification (IPC) or to both national classif	ication and IPC	·
B. FIELDS	SEARCHED		
Minimum d IPC 6	ocumentation searched (classification system followed by classifical C12N A01H	ation symbols)	
Documenta	tion searched other than minimum documentation to the extent that	such documents are included	in the fields searched
	lata base consulted during the international search (name of data b	ase and. where practical, sea	rch terms used)
	ENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the re	elevant passages	Relevant to claim No.
Α	WO 95 15381 A (CHAPELLE ALBERT D JOHNS HOPKINS (US)) 8 June 1995 see the whole document	E ;UNIV	1-5, 33-35
A	WO 97 37011 A (SETRATECH S A R L RHONA HARRIET (GB); LOUIS EDWARD 9 October 1997 see abstract see the whole document	;BORTS JOHN (GB)	1-35
A	WO 90 07576 A (SETRATECH) 12 Jul see the whole document and speci 5, line 16-31, examples 3-4		1,7-27, 31,32
		-/	
X Furth	er documents are listed in the continuation of box C.	X Patent family memb	pers are listed in annex.
'A" documer consider to filling da 'L' documer which is citation 'O' documer other m'P" documer	nt which may throw doubts on priority claim(s) or s cited to establish the publication date of another or other special reason (as specified) nt referring to an oral disclosure, use, exhibition or	or priority date and not a cited to understand the invention.  "X" document of particular recannot be considered in involve an inventive step.  "Y" document of particular recannot be considered to document is combined w	after the international filing date in conflict with the application but principle or theory underlying the levance; the claimed invention overlor cannot be considered to be when the document is taken alone levance; the claimed invention involve an inventive step when the with one or more other such docunities of the country of the claimed invention in being obvious to a person skilled same patent family
Date of the a	ctual completion of the international search	Date of mailing of the int	
16	April 1999	03/05/1999	
Name and m	aiting address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Ripswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nt. Fax: (+31-70) 340-3016	Authorized officer  Mateo Rose	11, A.M.

onal Application No C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT PCT/EP 98/06977 Citation of document, with indication where appropriate, of the relevant passages Relevant to claim No. Α WO 97 01634 A (ANGELETTI P IST RICHERCHE BIO ; JIRICNY JOSEF (IT); PALOMBO FABIO () 1.2 16 January 1997 see page 1, line 18 - page 2, line 28 see page 58 PROLLA T A ET AL: "MLH1, PMS1, AND MSH2 Α INTERACTIONS DURING THE INITATION OF DNA 1,2,4,5 MISMATCH REPAIR IN YEAST" SCIENCE. vol. 265, 19 August 1994, pages 1091-1093, cited in the application Α A. WATANABE ET AL., : "Genomic organization and expresssion of the human 1-6 MSH3 gene" GENOMICS. vol. 31, 1996, pages 311-318, XP002099967 see the whole document and specially Fig.3 Α ACHARYA S ET AL: "hMSH2 forms specific mispair-binding complexes with hMSH3 and 1-6 hMSH6" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, no. 93, October 1996, pages 13629- 13629, XP002080159 cited in the application sequences from this paper deposited in GenBank under AC= U61981 (hMSH3), and U54777 and U73732-7 (hMSH6). see the whole document Α K. LIU ET AL., : "characterization of the mouse Rep-3 gene: sequence similarities to 1-6 bacterial and yeast mismatch-repair proteins" GENE, vol. 147, 1994, pages 169-177, XP002099968 see the whole document and specially Figure 2. I. IACCARINO ET AL., : "MSH6, a Α Saccharomyces cerevisiae protein that 1-6 binds to mismatches as a heterodimer with MSH2" CURRENT BIOLOGY. vol. 6, no. 4, April 1996, pages 484-486, XP002099969 sequence is deposited at GenBank under AC= Z47746. see the whole document -/--

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

PCT/EP 98/06977

C.(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	PCI/EP 98/06977
Category '	Citation of document, with indication where appropriate, of the relevant passages	
	- sprophilis, or the leavant passages	Relevant to claim No.
Α	CORRADI A ET AL: "CDNA SEQUENCE, MAP, AND EXPRESSION OF THE MURINE HOMOLOG OF GTBP, ADNA MISMATCH REPAIR GENE" GENOMICS, vol. 36, no. 2, 1 September 1996, pages	1-6
	288-295, XP000613761 see the whole document and specially Fig.1	
P,X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,13 July 1998, XP002099368 HINXTON, GB AC= AJ007791. Arabidopsis thaliana mRNA for mismatch repair protein (MSH3) see abstract	1-6
P , X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,1 August 1998, XP002099369 HINXTON, GB AC= 065607. Arabidopsis thaliana. Putative mismacht DNA repair protein see abstract	1-5
P, X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 16 October 1997, XP002099371 HINXTON, GB  AC= AF009657. Arabidopsis thalian MutS homolog 6-like protein mRNA. see abstract  -& K.M. CULLIGAN AND J.B. HAYS: "DNA mismatch repair in plants" PLANT PHYSIOLOGY, vol. 115, 1997, pages 833-839, XP002099372 see the whole document	1-6
-	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,17 December 1998, XP002099373 HINXTON, GB AC= AJ131669. Triticum aestivum mRNA for MSH3 protein, MSH3 gene. see abstract	1-6
	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,12 October 1998, XP002099370 HINXTON, GB AC= AJ007792. Arabidopsis thaliana DNA mismatch repair protein, MSH6 gene. see abstract	1-6

G

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

Information on patent family members

PCT/EP 98/06977

Patent document		Publication			98/06977
Cited in search repo		aie	Patent fa membe	amily er(s)	Publication date
W0 9515381	A	08÷06-1995	EP 07: JP 956 US 569	30648 A 06509 T 93470 A 37443 A	11-09-1996 30-06-1997 02-12-1997 17-11-1998
WO 9737011	A	09-10-1997	NONE		
WO 9007576	A	12-07-1990	AT 12 AU 483 CA 200 DE 6892 DE 6892 EP 044 ES 207 IE 7	11793 A 27519 T 4390 A 6549 A 4174 D 4174 T 9923 A 7058 T 2469 B 3601 T	20-07-1990 15-09-1995 01-08-1990 26-06-1990 12-10-1995 14-03-1996 09-10-1991 16-11-1995 09-04-1997 02-07-1992
WO 9701634 	A	16-01-1997	IT RM950 AU 6241	0434 A 1296 A	27-12-1996 30-01-1997